

## **FIGURE 1**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG  
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT  
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT  
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG  
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG  
CTACCATTTATGTTTCGTTATAAGCAAGTTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA  
TTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCT  
CCAGAAAACAACCCCTTTTGTCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT  
TATATATGTTTGTTCAGACCATCCTTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTC  
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC  
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG  
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT  
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA  
TGGATTAACCCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA  
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT  
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT  
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT  
ATG

## **FIGURE 2**

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFI FSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV  
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG  
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW  
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR  
LLSRDI

**Important features:**

**Type II transmembrane domain:**

amino acids 13-33

**Other Transmembrane domains:**

amino acids 54-73, 94-113, 160-180, 122-141

**N-myristoylation sites:**

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

### FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC  
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC  
CTTACACTTCGCCATGAGTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG  
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG  
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG  
AGTATTGAATAGCAGCTCCCCTTATTTTCACTGGAAAATGAACCTGTGTGTAATCTGCTGATCCTGG  
TTTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA  
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC  
CATTTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG  
TGACTCTCATGGCTCTTCTTTCTGGATTTGGTGTCTCAACTGCCCATACACTTACATGCTTACTTC  
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT  
CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA  
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAATCTTACT  
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT  
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTT  
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT  
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT  
GAAGTTTTGGTCCCAACATTTCTTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT  
TGCTGATCACTCTTACCAAGTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTG  
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT  
AGAATACCGCACCATTAATCACTGAAGTCCTTGGAGAAGTGCAGTTCAACTTCTATCACCGTTGGTTTG  
ATGTGATCTTCTGGTCAGCGCTCTCTTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA  
GAGAAGCAAATGGCACCTTGAACCTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA  
GATATAAGAGGGGGGAAAAATGGAACAGGGCCTGACATTTTATAAACAACAAAATGCTATGTTAGC  
ATTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAG  
AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG  
TGTAGAGGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT  
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTTACATGGAAAAGGTTATAGCTTG  
CCTTGAGATTGACTCATTAAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG  
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

## **FIGURE 4**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV  
LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLEFSCLLWLTFMYFFWKL GDP  
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ  
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDAL EELSRQ  
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTD PVTRGIEI  
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGY  
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLV SALSSILFLYLAHKQAPEKQMAP

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
425-444

#### **N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

#### **Eukaryotic cobalamin-binding proteins**

amino acids 151-160

## FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC  
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT  
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT  
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT  
 GTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG  
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC  
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTTCGTGGTCATCTCTAG  
 GATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCCAGTGA  
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA  
 ATTATCACCACCAAGATCCCATATTCAACACTCAAACGCAACACAAACAGAATTTATTGT  
 CAGTGACAGTACCTACTCGGTGGCATCCCCTTACTCTACAATACCTGCCCCTACTACTCCTC  
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATGATTTGTGTACAGAAGTTTTTATG  
 GAACTAGCACCATTGTCTACAGAACTGAACCATTGTTGAAAATAAAGCAGCATTCAAGAATGA  
 AGCTGCTGGGTTTGGAGGTGCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTGGTGCTG  
 CAGCTGGTCTTGGATTTTGTCTATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT  
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAA  
 TGAGGAATCAAAGAAAACCTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC  
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTT  
 CATGCTCCTTACCCTGCCCCAGCTGGGGAATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCA  
 CCCTTGGTTTCTTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT  
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT  
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTCGAAAGTGCAAGGAC  
 CTAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTTG  
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCCCTTCTTCA  
 GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTAAGAGCAAAAGAAT  
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA  
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAAC  
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA  
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA  
 AAAATAAAAACCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT  
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC  
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGCACATAGTAG  
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT  
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTCTCTAATATGGCAACCATTATAACCTTAATT  
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC  
 ATTAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT  
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

## **FIGURE 6**

MARCFSLVLLLSIWTTTRLIVQGSRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG  
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN  
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTTPAPASTSIPR  
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFFGAAAGLGFCYVK  
RYVKAFFFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPESKSPSKTTVRCLEAEV

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

### **Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

### **Tyrosine kinase phosphorylation site.**

amino acids 79-88

### **N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

## FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCCGCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCCGGC  
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCCGGCCGAGCGCAACTCGGTCCAGTCGGGGCGG  
CGGCTGCGGGCGCAGAGCGGAGATGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGTGGCGG  
CGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGTCCAGTCAAGCCCCGGCCCG  
GCTCTCAGCTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTGT  
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA  
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCAGCTATCACAATGAGACCAACACAGAC  
ACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAATTACAAAGATAACCAACAACAGAC  
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC  
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGGCAGCTTCCAGTAC  
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA  
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA  
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC  
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC  
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC  
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG  
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCC  
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG  
CCGCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA  
TAGAAATAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCTTA  
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGCT  
CCCCAGGCTGTTTCTCAGGCTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAACTGCA  
GGAGCAGTTTGGCACCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG  
TTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAACCTGGCAAAAATG  
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC  
AGATGAAATGTTCTGTTCAACCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC  
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG  
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC  
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT  
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT  
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAGTAAA  
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT  
ATTGACACTGTCCCTCTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCA  
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC  
AAATCACTTAGCAGCACTGAAGACAATTATCAACCACGTGGAGAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG  
TTTTCAGGTGTGATGACTGTTGCCACCATGTATTATCCAGAGTCTTAAAGTTTAAAGTTGCA  
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA  
ATCAAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 8**

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL  
RSAVEEMEAEAAAASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFESE  
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRLCSTRDSECCGDQLCVWGHC  
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLPVCTPLPVEGELCHDPASRLDLITWELEPDG  
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE  
RSLTEEMALGEPAAAAAALLGGEI

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331

### **N-myristoylation site.**

amino acids 202-208, 217-223

### **Amidation site.**

amino acids 140-144



## FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC  
GGCCACCTTGTGAACCTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG  
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTTGGGGCTCTTCTGGACCCTT  
AACTGGGTACTGGCCCTGGGCAATGCGTCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT  
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC  
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG  
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT  
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA  
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC  
ATTGTCAGGGTGGTCGTCCTGGACAAAGTACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT  
GGTCGGAGGCGTGGGGGTCTGTCTTCTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG  
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT  
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT  
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA  
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG  
CCCTGATCCAGGACTGCACCCACCCCAACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT  
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT  
TTGAGAGGCTGAGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG  
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA  
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA  
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA  
AAGATTTTATTAAAGATATTTTGTAACTC

## **FIGURE 10**

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGGLFWTL  
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL  
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN  
IVRVVVLDKVTDLLFFGKLLVVGGVVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY  
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

### **Important features:**

#### **Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

#### **N-glycosylation sites.**

amino acids 16-20, 27-31, 289-293

#### **Hypothetical YBR002c family proteins.**

amino acids 276-288

#### **Ammonium transporters proteins.**

amino acids 204-231

#### **N-myristoylation sites.**

amino acids 60-66, 78-84

#### **Amidation site.**

amino acids 306-310

## FIGURE 11

GCCCCGCGCCCCGGCGCCGGCGCCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC  
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC  
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCTGCGGGTGTGGTGTCCA  
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC  
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG  
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA  
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCTGATCCTGGTGGGCCTC  
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGTTTCTACTTCGGCGTCTGTTGG  
CTCCTTCTCTTCATCCTCATCCAGTGGTGTCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT  
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC  
TTCTACTTGCTGTGATCGCGGCGTGGCGCTGATGTTTGTACTTACTGAGCCAGCGGCTGCCA  
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTCTGC  
CCAAGGTCCAGGACGCCCAGCCAACTCGGGTCTGCTGCAGGCCTCGGTATCACCTCTACACCATG  
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCCAGCT  
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG  
TGGGCCTCATCATCTTCTCTGTGCACCTCTTCATCAGTCTGCGTCTCTCAGACCACCGGCAGGTG  
AACAGCCTGATGCAGACCGAGGAGTGCCCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT  
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGCGTCACCTACAGTACTCCTTCTTCC  
ACTTCTGCCTGGTGTGGCCTCACTGCAGTCTATGATGACGCTCACCAACTGGTACAAGCCCCGGTGAG  
ACCCGGAAGATGATCAGCACGTGGACCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT  
CCTCTACCTGTGGACCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA  
CAGCCTGCCATCTGGTGCTCTGCCACCTGGTGCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC  
CCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCTGAGCCGGGC  
CTTCTAGTCGTAGTGCTTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC  
ACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA  
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCACAG  
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTTCTGGTCACGTCCCCCAGGGGACCCTGCCCCCTCCTG  
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAACAAGCCAGTGCGTGTAAAAAAA

## **FIGURE 12**

MGACLGACSLSCASCSCGSA PCILSCCPCASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL  
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ  
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSEFLFILIQLVLLIDFAHSWNQRWLKAE  
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV  
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI  
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY  
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD  
FS

### **Signal sequence:**

amino acids 1-20

### **Transmembrane domains:**

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

## FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA  
GGTTGGAAGAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAGAACG  
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA  
CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAGAAAGGCATATCTGATGTCAGGAGGAC  
TTTCTGTTTGTGTTGTACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG  
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT  
TTTGATATATTTCTTCTGGCAGTTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG  
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG  
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTTCATC  
CTTGCTGGATTGAGACGTGGTTCCTGGATTTCAGAGTTTACCTCAAGAAGCAGAAGAAGAAAA  
CAGACTCCTGATAGTTTCCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG  
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC  
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC  
CACTGCTGGCTTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC  
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTTCCTCTCT  
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCTTAATC  
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC  
AGAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT  
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTGTACAGTTTTAAAAAGTGAATAAAATCTG  
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTGTTTGCATTATGTGTATGGCCTGAAGTGTGGA  
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAATGTCACCAGACATTTGTATTATTT  
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC  
ACAAAATGACTTAAACCATTATATCATGTTTCTTTGCGTTTCAGCCAATTCAATTAAATGAA  
CTAAATTAAAAA

## **FIGURE 14**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF  
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT  
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER  
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-20

**Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

**N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

## FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG  
 CCGCGGCCCTGCCCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCCGGGACAGAAGATGTGCTCCAG  
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT  
 CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC  
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCCTTTGAGAACGGCATCACCATGCTCGACGC  
 AGGCAGCTTTGCCGGCTGCCGGGCTGCAGCTCCTGGACCTGTCACAGAACCAGATCGCCAGCC  
 TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG  
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA  
 CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC  
 AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC  
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGCGCTGCGGT  
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC  
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCTGACG  
 CGCCTGCGGCTGCGCGGCAACCCCGCATGCCCAGCTGCGGGCCGAGGACCTGCGCGCCTGGC  
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC  
 TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTCCCCCTGAGC  
 TGTTTGGCCCTGGGTGCGCGAGAGCCACGTCACTGCGCAGCCCTGAGGAGACGCGCTGCCA  
 CTTCCCGCCCCAAGAAGCTGCGCGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCCAG  
 CCACCACCACACAGCCACAGTGCCCAACGAGGCCCGTGGTGCAGGAGCCACAGCCTTGCTCT  
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC  
 CACTGCCCCACGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA  
 ATGGGGGCACATGCCACCTGGGGACACGGCACACCTGGCGTGCTTGTGCCCCGAAGGCTTCAGG  
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACAGTCACGCCGAG  
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC  
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG  
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCA  
 GCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTGTCATGCCCTTTGGGGCCCGGGCGGTGCCGGAGG  
 GCGAGGAGGCCTGCGGGGAGGCCCATACCCCCAGCCGTCCACTCCAACACGCCCCAGTCACC  
 CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC  
 GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG  
 ACAAGGGGAGGTGGGGCCAGGGCTGGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCTTGGAG  
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC  
 ACTCATGGGCTTCCAGGGCTTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA  
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC  
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT  
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC  
 CTAACGTCCCCAGAACCAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAATC  
 CCTGGGCACGGCGGGCCTGCCATGTGTGTGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC  
 TCCAGGCGGACCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC  
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC  
 TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT  
 GGGAAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAGACAAACGATGATATGAA  
 GGCTTTTTGTAGAAAAAATAAAGATGAAGTGTGAAA

## **FIGURE 16**

MCSRVP L L L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T  
M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y  
L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L D T A N V E  
A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L  
A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E  
T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P  
S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P  
V T P R P P R S L T L G I E P V S P T S I R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y  
T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V  
L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E  
C E V P L M G F P G P G L Q S P L H A K P Y I

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 579-599

#### **EGF-like domain cysteine pattern signature.**

amino acids 430-442

#### **Leucine zipper pattern.**

amino acids 197-219, 269-291

#### **N-glycosylation sites.**

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

#### **Tyrosine kinase phosphorylation sites.**

amino acids 124-131, 337-345

#### **N-myristoylation sites.**

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,  
594-600, 640-646



## FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG  
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCTCGG  
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA  
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA  
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAAGTGTACAG  
AAGATATCAGCTTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA  
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT  
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG  
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT  
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATGGAATGAAAATCCTTAATGGAAG  
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA  
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG  
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG  
CTTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTCTTGTATATTATACAT  
TTGGAGCTCTTGGGGGAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT  
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTCAGCTTTTCATGATC  
CAGATTTGCTTGTTAATTAAGACCAAATATTCAGTTGAACTTCCTTCAAATCTTGTTAATGGATAT  
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG  
TTTGGCTGATTGCCCCATAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG  
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA  
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC  
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAATGACTCCAGAGAGCTCTACTTTCTG  
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT  
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG  
AAGATTAACCTCATTTTTAATAAAATTATGTCTAAGATTAATAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 18**

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTTLTSDESVDHTTAGRVVAGQIFLDSESESEL  
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFFFLFDK  
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARKROMQEAEMMYQTGMKILNGSNKKSOKR  
EAYRYLQKAASMNHTKALERVSYALLFGDYL PQNIQAAREMF EKLTEEGSPKGQTALGFLYASGLGVN  
SSQAKALVYYTFGALGGNLI AHMVLVSRL

### **Important features:**

#### **Signal peptide:**

amino acids 1-21

#### **N-glycosylation sites.**

amino acids 195-199, 217-221, 272-276

#### **Tyrosine kinase phosphorylation site.**

amino acids 220-228

#### **N-myristoylation sites.**

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

#### **Glycosaminoglycan attachment site.**

amino acids 267-271

#### **Microbodies C-terminal targeting signal.**

amino acids 299-303

#### **Type II fibronectin collagen-binding domain protein.**

amino acids 127-169

#### **Fructose-bisphosphate aldolase class-II protein.**

amino acids 101-119

## FIGURE 19

AATTCAGATTTTAAGCCCATTCCTGCAGTGAATTTTCATGAAC TAGCAAGAGGACACCATCTTCTT  
GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG  
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG  
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACCTTTTGAT  
AAAAAGGGATTTCATGTAATCGCTGCCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA  
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG  
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGTCTCTGGGGTCTGATCAATAATGCTGGTGTT  
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA  
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG  
TTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA  
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC  
ATGCATTGAACCAGGATTGTTCAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC  
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAA  
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG  
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA  
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA  
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA  
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT  
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT  
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT  
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA  
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGT  
ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAA

## **FIGURE 20**

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTESG  
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY  
REPIEVNLFGLISVTLMMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK  
AFGVHVSCEPGLFKTNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD  
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 136-152

#### **N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

#### **Glycosaminoglycan attachment site.**

amino acids 39-42

#### **N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

## FIGURE 21

CTGAGGCGGCGGTAGCATGAGAGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGGC  
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTGGGGAAGTAAAA  
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA  
CATTAGAAATATATTCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG  
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT  
CGTCATTAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTTCAGGAGCATT  
TTCAAACCAAGACCTTGTTTTCTGCTATTAACCAAGTATAATAACAGAAAGCTGCTCTACTC  
ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT  
GCCAATCTGGGCATGTCTGAACAACCTGGGTATAAACTGTATCAGGTTCTGTATGTCCACTGG  
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG  
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAAGTG  
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA  
GAAAAGGAGAGGAGCACAGATTAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA  
ACATTTTTCTTTGTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTT  
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACACCACCTCTCGATGTAGT  
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC  
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA  
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT  
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC  
CTACATTTTGAATCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC  
ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCAATTTGTTTTTACTATGTTTACC  
TGTTTGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA  
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCCTAAGTTGTTGAGGGGAAGGCTTACACAG  
ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT  
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT  
GAGACCATGTCTATTAAAAATAAAATGGAAAAGCAAGAAATAGCCTTATTTTCAAATATGGAAA  
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTCCTTAAGTGATACTTTTTTAGAAGTA  
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT  
AAAATTTAAAAA

## **FIGURE 22**

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYITIDIQKYI  
PCYQLFSFYNSSGEVNEQALKKILSNVKNVVGWYKFRHSDQIMTFRERLLHKNLQEHFSNODL  
VFLLLTPSIITESCSTHRLEHSLYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV  
QTHSSKFFEEEDGSLKEVHKINEMYASLQEEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA  
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMMLKNRHSVSKSSCNYNHHLDDVVDNLTL  
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDLDQDKRSKANTGSSNQDKASKMSSPET  
DEEIEKMKGFGEYSRSPTE

### **Important features:**

#### **Signal peptide:**

amino acids 1-19

#### **N-glycosylation sites.**

amino acids 75-79, 322-326

#### **N-myristoylation site.**

amino acids 184-154

#### **Growth factor and cytokines receptors family.**

amino acids 134-150

## FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA  
GCAGCGCGCAGCGAACGCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCTGCCACCCTTCCTCTCCCTCC  
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCGAGGTCACCAGCCCGCGCTCT  
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCGCGCCTGGCACCGGGACCGTTGCCTGA  
CGCGAGGCCCAGCTCTACTTTTCCGCCCGCTCTCCTCCGCTGCTCGCTCTTCCACCACTCCAACCTCTTCTCCC  
TCCAGCTCCACTCGCTAGTCCCGACTCCGCCAGCCCTCGGCCGCTGCGTAGCGCCGCTTCCCGTCCGGTCCAAA  
GGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTGCGGTTGCCGCGCTTCTCTGCACCTGGCAGTGCTC  
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGTCTTTACGTGTCCAAAGGCTTC  
AACAGAACGATGCCCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT  
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG  
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTG  
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAAATCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG  
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG  
TTCCGCTTGGTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG  
CCCTTCGGAGATGTCCCTCGCAAATTGAAGTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGC  
TTAGCGGTGCGGGAGATGCTGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGATCCCATGCCCTGTTGAAG  
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCATGTTACAACACTGCTCAAACATCATGAGAGGC  
TGTTTGGCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA  
GAGGGTCTTTCAACATTGAATCGGTGATGATCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT  
AATAGTGTTCAAGTGCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT  
TCCATCTCTGAAAGTGCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT  
AGTTTGGACCGACTGGTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCCGAGCAAC  
GTTTGAACGATGAGAGGATGGCTGCAGGAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC  
CTGTTGCACTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAAACCAGAC  
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG  
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGCTCCTGGGGCA  
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCCTGGTTATGCAGAGAGAGTGAGATTAATTCTCAAACCTCTGAG  
AAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA  
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGTTTAAAGAGTGCTGACTTTGTTTTCTCATTCAAGTTTGGG  
AGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGAGGTACAGAA  
CTATAGTTAGTTGTGCATTGTGATTTTATCACTCTATTATTTGTTGTATGTTTTTCTCATTTCGTTTGTGGGTT  
TTTTTTTCCAACGTGATCTCGCCTTGTTCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATT  
TCTGAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTATTAAAGAAAAAGCCCAAAAGC

## **FIGURE 24**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST  
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH  
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRVLVNSQYHFTDEYLECVSKYTE  
QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL  
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN  
MQDNSVQVSQKVFQCGPPKPLPAGRISRSESISAFSARFRPHHPEERPTTAAGTSLDRLVTDVK  
EKLKQAKKFWSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT  
KPDILILRQIMALRVMTSKMKNAYNNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG  
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **ATP/GTP-binding site motif A (P-loop).**

amino acids 515-524

#### **N-glycosylation site.**

amino acids 514-518

#### **Glycosaminoglycan attachment sites.**

amino acids 494-498, 498-502

#### **N-myristoylation sites.**

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

#### **Glypicans proteins.**

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506



## FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC  
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT  
TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA  
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC  
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG  
AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCTGTGATCATT  
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA  
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG  
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC  
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCAGTGCTCTCAAAAAGCA  
TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT  
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT  
AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT  
AAATGTCAAAAAAAAAAAAAAAAAA

## **FIGURE 26**

MKVLISL L L L L L L L L L L P L M L M S M V S S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R K F M  
TVSGLPKKQCPDHF KGNVKKTRHQRHHRKPNKHSRACQQLKQCQLRSFALPL

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-myristoylation sites.**

amino acids 27-33, 46-52

## FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG  
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTGCTTCTTACCCTGCC  
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCCTACCTGA  
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG  
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC  
CAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA  
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT  
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTCTGTG  
CTCTGTGCAGAGCCCCAAGGAAGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC  
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC  
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA  
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG  
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTGTCTCC  
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA  
GCAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC  
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG  
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC  
CCAATGTTGTCCCTTTCCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC  
CCATGCGTCTCTAGGAACGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC  
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG  
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA  
ACCACG

## **FIGURE 28**

MDILVPLLQLLVLLLTLP LHL MALLGCWQPLCKSYFPYLM AVLTPKSNRKMESKKRELF S QIKGL  
TGASGKVALLELGC GTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM  
RQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRVLRPGGV LFFWEHVAEPYGSWAFMWQQVF EPTW  
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWL PVGPHIMGKAVKQSF PSSKALICSFPSL  
QLEQATHQPIYLPLRGT

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Leucine zipper pattern.**

amino acids 10-32

#### **N-myristoylation sites.**

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

## **FIGURE 29**

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT  
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG  
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT  
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA  
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA  
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTCATTAAGAAATT  
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC  
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

### **FIGURE 30**

MLLLTLLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG  
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

bioRxiv preprint doi: <https://doi.org/10.1101/2021.04.28.440000>; this version posted May 3, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

## FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT  
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT  
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGAATCATGTCGG  
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC  
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT  
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA  
CAGAAAGGGGAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG  
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC  
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA  
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCCAGGTT  
ATGGAAGGCGGCCAAGTGAATATAAGCCCCCTTCGGGCATTCCGTACATGTGGTCGTACCATT  
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG  
TGGTTACTTGTTATTTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTCGTCTCTC  
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG  
GATTCCGAGAAATCATTGTGATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT  
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC  
AACCAGAATGCATATACTACAACGCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC  
ATTCAAAATCTTGTTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA  
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC  
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC  
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC  
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA  
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA  
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTAGGTATCTGTACCTGGAAAACATT  
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT  
TTAAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

## **FIGURE 32**

MSGRDTILGLCILALALSLAMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYDTNDLSIE  
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA  
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA  
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG  
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD  
FIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVVHSLVVFETVLDALFLCFA  
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

### **Important features:**

#### **Signal peptide:**

amino acids 1-20

#### **Putative transmembrane domains:**

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

#### **N-glycosylation sites.**

amino acids 204-208, 295-299, 313-317

#### **N-myristoylation sites.**

amino acids 147-153, 178-184, 196-202, 296-275, 342-348



## FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTTAGA  
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT  
 GAAAATTTTTTGAAGAAAAAATTGCCTTCTTCAAACAAGGGTGTCTTCTGATATTTATGAGGAC  
 TGTGTCTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGTGCTGGTACTGGAGTAC  
 ATTCAAACAAGAAACGGCAAGAAGATTAAAGGCCCAAGTTCTGTGCCTCAGATCAACTGC  
 GATGTCAAAGCCGAAAGATCATCGATCCTGAGTTTCTGTTGAAATGTCCAGCAGGATGCCAAGA  
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG  
 TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT  
 GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT  
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATACTCATCAT  
 CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG  
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC  
 CACCTTGCCAAAGGCCATCCCCTTCTGTGCTTCTACCACAGCATCCCCAGACCACAATCAGTGG  
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC  
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCCTCCAGAAACCTGTTGGAGC  
 GGATGTCAAGCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC  
 TGGGAGATCCAACTGCAAAATTGACTTGTGCTTTTTAATTGATGGGAGCACCAGCATTGGCAAA  
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC  
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCCTGCTACTCACTTTAACCCTCAAGACAC  
 ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT  
 GTAGGTTCGGCCATCTCCTTTGTGACCAAGAATCTTTTCCAAAGCCAATGGAACAGAAAGCGG  
 GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
 GACTTGCGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG  
 AAGCAGTATGTGGTGGAGCCCAACTTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC  
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG  
 AACTGACCGCCTGGCCTGCAGCAAGACCTGCTTGAAGTGGCTGACATTGGCTTCGTCATCGAC  
 GGCTCCAGCAGTGTGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA  
 AGAGTTTGAGATTTCCGACACGGACACGGCATCGGGCCGTGCAGTACACCTACGAACAGCGGC  
 TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC  
 TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA  
 GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCTTACGACGACGTCC  
 GGATCCAGCCATGGCTGCCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCCTGGGCT  
 GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA  
 GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC  
 AGCCTCGGAACGAATTTCAGAGCAGGCAGAGCACCAGCAAGTGTGCTTTACTAACTGACGTGTT  
 GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC  
 AAATGTCTGTATTATTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTTGGAGTTACAAAGA  
 TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGCTGGAGATTTTACAT  
 TTTGACAATTGTTTTCAAATAAATGTTCCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG  
 AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACTCTGTAACCTCAGCAAGTTTCAT  
 TTTTGTGATGACAATGTAGGAATTGCTGAATTAAATGTTTAGAAGGATGAAAAATAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 34**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG  
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR  
ESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA  
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQRQDPSGAAFQKP  
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI  
GPAGPLMGVVQYGDNPATHENLKTHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN  
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG  
FYSLHVQSWFGLHKTLLQPLVKRVCDTDRACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN  
LTKEFEISDTRIGAVQYTYEQRLFEGFDKYSSKPDILNAIKRVGYWGGTSTGAAINFALEQL  
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHDSFF  
VDEFDNLHQYVPRIIQNICTEFNSQPRN

### **Important features:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 181-200

#### **N-glycosylation sites.**

amino acids 390-394, 520-524

#### **N-myristoylation sites.**

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,  
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

#### **Amidation site.**

amino acids 304-308

## FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC  
CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT  
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC  
GGTTGACCTGTTGGCTCTTCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCTGGAGTGA  
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCGAGTA  
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTCTTTGTGATGTGGTAC  
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA  
CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTC  
TGGTGACCTCCCACTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTAAAAAAGTCT  
TGGTGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC  
ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA  
ACCTGACCTTGAAAACCATTTATGGCATTCAAGTGGGTAAGTGAAGTATCTTTTAAACCTAAACCACTCAGA  
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA  
GAAGTTTTTCACAGGTTATCCTCTAATTGATAATTATTCTATAGAGGATTTTACCAAAAACCCATATTT  
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCCAGAGAT  
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAGGATCAAGTTGAAGATGTTTATGTCCGGAT  
CTGTTTGAATTTATTAAGTGAACATTCATATCCAGAAGACACAAATCTTTCTTTCTATATAGAATCC  
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCAAGGAGATCATCACTTTT  
TGGCAGGTGATGCTAAGGAACACCACATGCCATTATTAAGTTTACATTCTACAAAAGCCTAGAAGGACAG  
GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAATTCATGGGGAGGTCAAGTGTGCTGGCTT  
ACACTGAACTGAACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG  
CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG  
ACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCA  
CTAGGCTGTAAAAACAAAACATGTAGAGTTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA  
TATCTTATGTGGATTACCAATTTAAAAATATATGAGTTCTGTGTCAAAAACTCTTCACTGAAGTTATA  
CTGAACAAAATTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGAGTATTTACAGTTATT  
ATTATTTAAAAATTACTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG  
TGAATCATTTCTTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA  
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT  
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

## **FIGURE 36**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF  
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLA  
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL  
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV  
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR  
NTTCHY

**Important features:**

**Type II transmembrane domain:**

amino acids 20-39

**N-glycosylation sites.**

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

**Glycosaminoglycan attachment site.**

amino acids 239-243

**Ly-6 / u-PAR domain proteins.**

amino acids 23-37

**N-myristoylation site.**

amino acids 271-277

## FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC  
TCGTCCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG  
AATATGATCAGATTGAGTGCGTCTGCCCCGAAAGAGGGAAGTCGTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA  
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACCTGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT  
ACCTTGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCCGATGTG  
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCCATAATGCTCACTGTGAATGGACCATTCATGC  
TAAACCTGGGTTTGTATCCAACTAAGATTTGTTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG  
GTTTCGTGATGGAGACAACCGGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG  
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTATGAGGAGATCACAGC  
ATGCTCCTCATCCCTTGTTTCCATGACGGCACGTGCGTCTTGACAGGCTGGATCTTACAAGTGTGCTGCTGGCAGGC  
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTGGGGGGCCAGTCAATGGGTACCGAGAAAA  
TAACAGGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCGTGGTGTCTTTCTTTTGAACACTCCTATGT  
TCTTAGTGGCAATGAGAAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA  
GAACCAAGATTTTCAAGCTGGTGAGAAAGGAGAGTTCTTCCGATGCAGGTTTCAAGGAGAGACACCATACACAGCTAT  
ACTCAGCGGCTTCAGCAAGCAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG  
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT  
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA  
CCCAGGGTTGCGCTGGCCGTGGCAGGCAGCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC  
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTTACTGACCTGGGGAAG  
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATACCGGGATGAGAAGACCATCC  
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCTGAAGCT  
CCTAGACAAGGCCGATCAGCACCCGAGTCCAGCCCATCTGCTCGCTGCCAGTCGGGATCTCAGCACTTCTTCCAGGAG  
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG  
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT  
CTGTGCCAGCTGGGAACCCACTGCCCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGCTCTCCCGGA  
CGAGCATCTCTGAGCCAGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA  
CTGCCCTTACCAAGGTGCTGCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG  
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCCTGTGAATTGGCT  
GTGCCAGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGTTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA  
CTACTAGGACAGCCAATTGGAAGATGCCAGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACCATATACAAAACCTCTCCA  
CTCCACTGACCTGGTGGTCTTCCCCAACTTTTCAATTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG  
GCCCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCTGTGGGACAGCCAGGCGAGCAGAGCTGGGATGTGGTGCATGCTT  
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTCTCCCCATCTCTGTACACATTTTAAATAAATAAGGGTTGGCTTCT  
GAACTACAAAAA  
AAAAA

## **FIGURE 38**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT  
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRCGQ  
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII  
KRVCGNERPAPIQSIGSSSLHVLFHSDGSKNFDGFAIYEEITACSSSPCFHDGTCVLDKAGSYKC  
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE  
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPQMVSRETPHLQLYSAAFSKQKLQSAPTK  
KPALPFGDLPMPGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP  
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCSGALVNERTVVAAHCVTDLGKVTMIKTADL  
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR  
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVVSVVDSLLCEEQHEDHGIPVSVTDNMFCA  
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLGLVSWSYDKTCSHRLSTAFTKVLPFKDWI  
ERNMK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-23

#### **EGF-like domain cysteine pattern signature.**

amino acids 260-272

#### **N-glycosylation sites.**

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

#### **N-myristoylation sites.**

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,  
474-480, 491-497, 638-644, 666-672

#### **Amidation site.**

amino acids 56-60

#### **Serine proteases, trypsin family.**

amino acids 489-506

#### **CUB domain proteins profile.**

amino acids 150-167

## FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTTAATC  
TGAAGGTCTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG  
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC  
TGGTCTTGCCCTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTCAGAACTCTGTA  
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC  
TCACAGCCACGGCTCCCTCCCAAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA  
ACCTGTCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGGAGCAACCGAACTA  
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA  
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAAATTCTGAAAACACCACTGCCCTGAAG  
TCTTTCCAAGGTTGTACCACCTGATTCCAGATGTTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA  
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTATCGTGATGGGG  
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC  
ACAACTACGCTGTGCTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGTGACTGTGATGCGTGAACAGAAGTTCGCCA  
GCAGGAACAATGGACAGGCCCGGATGCCTACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC  
CCGAGGAGCAGCTTGAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTCAATGTGCTGGATGGCG  
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA  
GCCAGAAAAGTGGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCTGTCCCGCCAGGTTCCGGCAGC  
GGAGCCCTGACATCTTTGAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA  
CTCCCAAGCCCTCCATCCTACATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG  
GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCGGAGGAG  
TCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTGAATGTGGATGGGTCGAACTGACAGAGGTCAGCC  
GGAGTGAGGCAGTGGCATTATTGAAAAGAATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC  
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCT  
GGGTGATGTGGCTGGAATTACCACGGTGCTTGATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAGTC  
TGGGCTTCTGCATTGTAGGAGGTATGAAGAATAAATGGAACAAACCTTTTTTCATCAATCCATTGTTGAAGGAA  
CACCAGCATAAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA  
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT  
TTTTATAGAATCAATGATGGGTGAGAGGAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC  
TTGTCAGTTTTTATATTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT  
ACACCTCAGAAAAATGATTCAAAAAAATTAACCTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC  
AACATTGTTTATATTTTTCTATTCAATAAAAAGCCCTAAACAACTAAAATGATTGATTTGTATACCCCACTGAATT  
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGTACATTATGGCCATTTTTAATTTACAGCT  
AAAAATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAACAAGAATAAATATTTTTTCAGAAGTTAAA

## **FIGURE 40**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS  
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR  
ALSVLRRTKSGSAVANHADQGRENTTAPFVPRLYHLIPDGEITSIKINRVPSESLSIRLV  
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNMGDISNVPHNYAVRLLRQPCQVLWLTVM  
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDGGVAYRHG  
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRSPDIFQEAGWNSNGSWSPG  
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR  
IKTGDILLNVVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNNHMAPP  
SDWSPSWVMWLELPRCLYNCKDIVLRRTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG  
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

### **Important features:**

#### **Signal peptide:**

amino acids 1-15

#### **N-glycosylation sites.**

amino acids 108-112, 157-161, 289-293, 384-388

#### **Tyrosine kinase phosphorylation sites.**

amino acids 433-441, 492-500

#### **N-myristoylation sites.**

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,  
467-473, 603-609



## FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT  
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT  
TTCCACCTTTCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGTTGGGTGGGC  
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCC  
ATAAGACCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT  
GACAACTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC  
TTTGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG  
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA  
CATCTGCATCCCTTCTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG  
TAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG  
ACTGCTTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG  
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGATATTTTGG  
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGAT  
GGGGAGGCCAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG  
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG  
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT  
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTCTGGTTTGGTGCATGACCCC  
TGGATCTTTTGGTGATGTTTGGGAAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAA  
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATGTGTGAGCTGAATTTTTCTTTTGTATTTTCT  
TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT  
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA  
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTATTTAAATTTGAAGTAATATATTATGGGAT  
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA  
AGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT  
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA  
GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC  
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT  
TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT  
ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

## **FIGURE 42**

MGFNLT FHL SYKFRLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN  
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH  
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV  
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWWGED  
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV  
SVEHNPLYINITVDFWFGA

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

### **FIGURE 43**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG  
CCGCTTGCCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCAACAG  
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT  
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC  
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCTCCC  
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTTGTTT  
TCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI  
CIFCCGCCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

## FIGURE 45

GTGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC  
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT  
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG  
GACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA  
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAG  
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA  
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA  
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT  
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC  
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCTGTCAGCAGAACT  
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG  
GTCCTCCTGTGTCTCCTGTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG  
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG  
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT  
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA  
TCTTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCA

## **FIGURE 46**

MAGSPTCLTLIYILWQLTGSAAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP  
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK  
PKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC  
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE  
EKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT  
PRLFAYENVI

### **Important features:**

#### **Signal peptide:**

amino acids 1-22

#### **Transmembrane domain:**

amino acids 224-250

#### **Leucine zipper pattern.**

amino acids 229-251

#### **N-glycosylation sites.**

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,  
291-295

## FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG  
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT  
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT  
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG  
CTGCAACAACAGAACTGGAATGTTTCTTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC  
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCTCTCATGTGTAATTCTCCAAGC  
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA  
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG  
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG  
CTTATCCACTTCTCAGTATTTTtaggtctattgcttggtggaattctggaggtcctggttgggct  
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG  
TGTAGTTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

## **FIGURE 48**

MTCCEGWTS CNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGLIIGAGLMAIPA  
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL  
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSSEENKHRLIHFSVFLGL  
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

**Important features:**

**Transmembrane domains:**

amino acids 10-31 (type II), 50-72, 87-110, 191-213

**N-glycosylation sites.**

amino acids 80-84, 132-136, 148-152, 163-167

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 223-227

**N-myristoylation sites.**

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 207-218

**TNFR/NGFR family cysteine-rich region protein.**

amino acids 4-12



## **FIGURE 49**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA  
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG  
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTGAACTGTGA  
CATCGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC  
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC  
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG  
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA  
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC  
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCTT  
TTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

## **FIGURE 50**

MERVTIALALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS  
SQKQHSPVPEKAIPILITPGSATTC

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 36-59

**N-myristoylation sites.**

amino acids 41-47, 45-51, 84-90

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.**

amino acids 54-67

## FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG  
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG  
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG  
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG  
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA  
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG  
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC  
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA  
TTGGCAGACAGGCAGAAGATGTCTTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG  
GTGCCTGGCCACAGTGGTGTCTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG  
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG  
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA  
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGAG  
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA  
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC  
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG  
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT  
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA  
CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGAATCTGGGATTCAGGG  
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG  
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT  
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA  
GAATTTTAAATCCAAGCTGGGTTTTCATCAACTGGGATGCCATAACAAGGACCAGAGAAGCTCTC  
GCATCCCGTGAACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCAGCTCCCTCCTTAA  
AACACCACCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[illegible]

Signal peptide:

N-glycosylation site.

Glycosaminoglycan attachment site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Cell attachment sequence.

APP ID=10063732

## FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG  
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT  
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTC  
CACAGCCCCAAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC  
TTGAAGGACTCGACCCAGATGTCGGCCACCTATCCCAGGGCTTTACGGTATGGCTGGGTCCCAT  
CATCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA  
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG  
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT  
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG  
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA  
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT  
CTTGAGCTCAGTGCCCTTG TAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTCTGT  
ATTACCTCTCCCATGACGGGCGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC  
GCTGTTCATCCGGGAGCGGCGTCGCACCTCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA  
AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG  
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG  
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC  
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC  
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG  
CCTCATCGATATTATAGGGGTCCATCACAACCCAACCTGTGTGGCCGGATCCTGAGGTCTACGACC  
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA  
GGGCCCAGGAACATGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGAT  
GCTGCTGCACTTCCGGTTCCTGCCAGACCACCTGAGCCCCGAGGAAGCTGGAATTGATCATGC  
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC  
CCATCCACCTGTTTTTTTGAGATTGTCATGAATAAACGGTGTGTCAA

## **FIGURE 54**

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPQPPKRNWFWGHLG  
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP  
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI  
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC  
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDLFDVLLLSKDEDGKALSDEDIRAEADTF  
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR  
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP  
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN  
VGLQ

**Important features:**

**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 461-471

**N-glycosylation sites.**

amino acids 112-116, 168-172

## FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA  
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT  
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG  
TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTTTGCCGTGTGTCTTGC  
ATAATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT  
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGACGAATGTGTTGCTT  
GTGATTCGAACATTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA  
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTTCCTCATGTACCTGTTTCCTC  
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 56**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR  
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCAGCCAGCAGGCGGCCAACATGCTCTGTCTGTGCCTG  
 TACGTGCCGGTCACTCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC  
 ATTTTCAAGCTCAGTGTCTTTCATCCCTCCCGAGGAATTCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCT  
 GGAGATAAGGACCTTGATGGGCGAGTACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG  
 CTGGTGTGTTAAGATTTTGGACAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG  
 GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAGCATGGATAAAAAACGGCACGATGACCATCGACTGG  
 AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACG  
 ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA  
 CACCTGGTGGCAGGAGGTGGGGCAGGGGCGGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG  
 CAGGTCCATGCCTCCCGCAGCAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGGCCAGG  
 TCACTCTGGCGGGCAATGGCATCAACGTCCTCAAAATTGCCCGCAATCAGCCATCAAATTATGGCCTATGAGCAG  
 ATCAAGCGCCTTGTGGTAGTGACAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCCCTGGCAGGGGCC  
 ATCGCCAGAGCAGCATCTAGCCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA  
 ATGCTGGACTGCCCGAGGAGTCTGGCCAGAGAGGGGTGGCGGCTTCTACAAAGGCTATGTCCCCAACATGCTG  
 GGCATCATCCCTATGCCGGCATCGACCTTGCACTACAGAGCGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTG  
 AACAGCGCGGACCCCGGCGTGTGTTGCTCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC  
 CCCCCTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCTC  
 TTCAAACATATCTGCGGACCGAGGGGGCCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGTCAATCCCA  
 GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCTGAAGATCACCTTGGGCGTGCAGTGCAGGGTGGAGGGGAGGGG  
 CGCCCGGCGAGTGGACTCGCTGATCTTGGCCCGCAGCCTGGGGGTGTGAGCCATCTCATCTGTGAATGTGCCAACACT  
 AAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGAGGAGGCTGGGAGAGCTGGCAGGCCCCAGGGCTT  
 GTCCTGCTGACCCAGCAGACCTCCTGTTGGTTCAGCGAAGACCACAGGCATTCCCTAGGGTCCAGGGTCAAGCAG  
 CTCGGGCTCACATGTGTAAGGACAGGACATTTCTGCACTGCTGCCAATAGTGAGCTTGGAGCCTGGAGGCGCGCT  
 TAGTCTCTCCATTTACCCCTTGACCCAGCTGTTGGCCAGGCCCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC  
 CCTCTGTGCTGCCTGCTGTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCACCCCTCGTCCAATCCC  
 ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCTGACTTCCCAACCTACAGCATGACGCCAATTGGC  
 TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCTGCTGATGGCTGGGGCTCTCGGGCATG  
 CTTGGGAGTGCAGGGGCTCGGGCTGCTGGCTGGCTGCACAGAAGGCAAGTGTGGGCTCATGGTGTCTGAGCT  
 GGCTTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCCACTGTGGCATGAGGGCAGTGGAGCA  
 CCATGTTTGGGGCGAAGGGCAGAGCGTTTGTGTCTTGGGGAGGGAAGGAAAGGTGTTGGAGGCTTAATTATGG  
 ACTGTTGGGAAAAGGTTTGTCCAGAAGGACAAGCCGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG  
 GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACCCAGCAGGGGCGCAGC  
 GGGACAGCCCCACATTCCACTTGTGCTACTGCTTGGAACTTATTTATTTGTATTTATTTGAACAGATTATGCTCCT  
 AACTATTTTATAGATTTGTTTAAATTAATAGCTTGTCTATTTCAAGTTTATTTTATTCATATTTATGTTTATGTT  
 GATTGACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGAGAAGGGGGGCTTGGGCGCTGCAGTCACATCT  
 GTCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAGGCAGCCCTGGCTCCTTTCTTTTGGCAG  
 GTTGGGGAAGGGCTTGCCCCAGCCTTAGGATTTTCAAGGTTTGAAGTGGGGCGTGGAGAGAGGGAGGAACCTCAAT  
 AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTCTGAATGTCAAGGCAG  
 TGAGTGCCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCCAGC  
 CTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAATGACCTGA  
 TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTGTTATATATGAACATATAACTGGAGTCGTCAAAAAG  
 CAAATTAAGAAAGAAATGGACGTTAGAAGTTGTCAATTAAGCAGCCTTCAATAAAGTTGTTTCAAGCTGAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 58**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG  
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG  
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG  
AGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRNGINVLKIAPESAIAK  
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR  
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC  
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV  
VYENLKITLGVSQR

### **Important features:**

#### **Signal peptide:**

amino acids 1-16

#### **Putative transmembrane domains:**

amino acids 284-304, 339-360, 376-394

#### **Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

#### **N-glycosylation sites.**

amino acids 129-133, 169-173

#### **Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140

## FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC  
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGGATCATCATTATTCTGGCTGGAGCAATTG  
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCCTCAGCT  
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTTCTGATAT  
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG  
AGCTGTGCGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTGTTGCTGATCAAGTGATAGTT  
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT  
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG  
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCAGGTGTGAGGCTCCCCGATGGTTCCCC  
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC  
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA  
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG  
ACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT  
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT  
GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC  
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG  
AGCAAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT  
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTGATGTGAACTAGACAAGTGTGTTAAGA  
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT  
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC  
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC  
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG  
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC  
TCTTCCCACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTAGCATAAACAGAGCAGT  
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 6o**

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS  
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKKNVQLTDAGTYKC  
YIITSKKGKNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS  
NTSFELNSENVTMKVVSVLNVTINNTYSCMIENDIAKATGDIKVTESI KRRSHLQLLNSKASL  
CVSSFFAISWALLPLSPYMLK

### **Important features:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 258-281

#### **N-glycosylation sites.**

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,  
220-224

#### **N-myristoylation sites.**

amino acids 52-58, 126-132, 188-194

## FIGURE 61

TGACGTCAGAATCACCAATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG  
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT  
GGGCTACCCCTGGTGGTGGTTATGGGGGTCCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG  
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAACTCCAGGAGGACCATATG  
GCGGTGCAGTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT  
GGGCTTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC  
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA  
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA  
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAACCTCTT  
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC  
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCCTCT  
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA  
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA  
CCATGACAGCTTCTCGGATGCTATTGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT  
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCCTGTCCCTCTAGAAGAAC  
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTTGCATCATAGCCACCA  
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA  
ATGTCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCCTTGTAATGG  
AGTTAGTGTCCAGTCAGCTGAGCTCCACCTGATGCCAGTGGTGAAGTGTTCATCGGCCTGTTACC  
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAAT  
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCACTGGCTGGATTCTGCCACACCCATAAAT  
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT  
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT  
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTC  
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

## **FIGURE 62**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG  
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH  
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNL FQQYDR  
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK  
DTAVQGNIRLSFEDFVTMTASRML

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-  
66, 70-75, 78-83, 83-88, 87-92, 110-115

## FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC  
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC  
AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCTCTAACTCTGAACATCCGCAGCCCGC  
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG  
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG  
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA  
AGCGCTGCCTGAAGAACTCTTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT  
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC  
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCACTGGGAGCCGGGGGAAAAATCCTTTCCCAACG  
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG  
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA  
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG  
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG  
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCCTCCTGGAGTTCTCCGC  
CCTCCTGGCTCTTCTTGGAAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG  
GGGCTTAGAGCACGATAGAGGGAAACCCAAACATTGGGAGTTAGAGTCCTGCTCCCGCCCCTTGCTG  
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC  
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 64**

MQGRVAGSCAPLGILLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL  
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAAEADRLGEA  
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP  
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTTRMPHPEGIWGINNQPPGTSGWNINRYPGGS  
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

**Important features of the protein:**

**Signal peptide:**

amino acids 1-26

**Casein kinase II phosphorylation sites.**

amino acids 56-59, 155-158

**N-myristoylation sites.**

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,  
259-264, 269-274, 270-275, 280-285, 281-286, 305-310



## **FIGURE 65**

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC  
TGGGCTGCCCCCTTGTCCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA  
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA  
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCCTCCACCATGC  
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC  
GGGCTTTTGGGCCCGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCACCCTC  
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

## **FIGURE 66**

MGSGPLVLVLLTLLGSSHGTPGMTLQLKLKESFLTNSSEYESSFLELLEKLCLLLHLPSTSVTL  
HHARSQHHVVCNT

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 37-41

**N-myristoylation sites.**

amino acids 15-21, 19-25, 60-66

## **FIGURE 67**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC  
CAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC  
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG  
CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC  
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG  
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG  
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA  
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAAGGGCT  
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA  
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA  
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTTATATATTTATATAAAATTAG  
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MANPGLGLLLLALGLPFLLRWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVFS  
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

### **Important features:**

#### **Signal peptide:**

amino acids 1-19

#### **Transmembrane domain:**

amino acids 56-80

#### **N-glycosylation site.**

amino acids 36-40

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 86-90

#### **Tyrosine kinase phosphorylation site.**

amino acids 86-94

#### **N-myristoylation sites.**

amino acids 7-13, 26-32

# FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCACCAG  
TCAAATACTTCCCTTATTAAAGCTGAATAATAATGGCTTTGAAGATATTGTATTGTTATAGATCCTAGTGTGCCAGAA  
GATGAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA  
TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATGAA  
AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATAACCAAGCAGTTTACA  
GAATGTGGAGAGAAAAGCGAATACATTCACTTCACCCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA  
CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTTC  
TACCGTGTAAAGTCAAAAAAATCGAAGCAACAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT  
CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACGTATGGAAAAGATTGTCAATTCCTT  
CCTGATAAAGTACAAACAGAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAA  
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT  
TCTGAGGATTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA  
AGAAATGTGTGCTTAGTTCTTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCTAAATCGAATGAATCAAGCAGCA  
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACCTTTGATAGTACTGCCACTATTGTA  
AATAAGCTAATCCAAATAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA  
ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGGTGATTGGAGAGCTACATTCCCAACTCGATGGATCCGAAGTA  
CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTGAACAAAAGTGGGGCCATTGTTTCA  
TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCAATTTTATGTT  
TCAGATGAAGCTCAGAACAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCGAAG  
TCCCTTCAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCTGGATGAACGACACTGTCTAATTTGATAGTACA  
GTGGGAAAGGACAGCTTCTTCTCATCATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCAGTGGGAACA  
ATAATGGAAAATTCACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAATGCAAAAGTGGGCACT  
TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAAACATTAACATATTACAGTAACCTTCTCGAGCAGCAAAATCTTCT  
GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCGAGCCCAATGATTGTTTACGCAGAA  
ATTTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTATTGAATCACAGAATGGACATACAGAAGTT  
TTGGAACTTTTGGATAATGGTGCAGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT  
ACAGAAAATGGCAGATATAGCTTAAAAGTTGGGGCTCATGGAGGAGCAAAACACTGCCAGGCTAAAATTACGGCTCCA  
CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAACCCGCCAAGACCTGAAATTTGAT  
GAGGATACTCAGACCACCTTGGAGGATTTTCCAGCCGACAGCATCCGGAGGTGCATTTGTGGTATCACAGTCCCAAGC  
CTTCCCTTGGCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCACAGTTTATGAGGATAAGATTATTCTT  
ACATGGACAGCACCAGGAGATAATTTTGTGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTCAAGTATTCTT  
GATCTAAGAGACAGTTTTTGTGATGCTCTTCAAGTAAATACTACTGATCTGTCCACAAAGGAGGCCAACTCCAAGGAA  
AGCTTTGCAATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTTATTGCCATTAAAGTATAGATAAAA  
AGCAATTTGACATCAAAAGTATCCAACTTGCACAAGTAACCTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGAT  
CCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTTCTGGAGTTAATATTCTACGCTGGTATTG  
TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTGGAACCTTAACGAAGAAAAAATCTTC  
AAGTAGACCTAGAAGAGAGTTTAAAAAACAAAACAAATGTAAGTAAAGGATATTTCTGAATCTTAAATTTATCCCAT  
GTGTGATCATAACTCATAAAAATAATTTTAAAGTGTGCGAAAAGGATACTTTGATTAAATAAAAACTCATGGATA  
TGTA AAAACTGTCAAGATTAAATTTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAG  
ATCCTTTTTTCTACTGATACCTGGTGTATATTATTGATGCAACAGTTTTCTGAAATGATATTTCAAATTTGCATCAA  
GAAATTTAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAATAAACCAACTTTGGA AAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

10063732

## **FIGURE 70**

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE  
ATEKRFFFKNVSILIPENWKENPOYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY  
IHFTPDLLLGGKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN  
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP  
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRNLN  
MNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK  
YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG  
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL  
ITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR  
AANSSVPPITVNAKMNDVNSFFSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA  
GADSEKNDGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP  
RPEIDEDTQTTLDEFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWAPGDN  
FDVGKVQRYIIIRISASILDLRDSFDDALQVNTDLSPEANSKESFAFKPENISEENATHIFIAI  
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVI  
VNFILSTTI

### **Signal peptide:**

amino acids 1-21

### **Putative transmembrane domains:**

amino acids 284-300, 617-633

### **Leucine zipper pattern.**

amino acids 469-491, 476-498

### **N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

# FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAGACCGGGAAAGACCATACGTCCCCGGGCGAGGGGTGA  
CAACAGGTGTGATCTTTTGTGCTGTGGTGCCTTCTTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA  
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCACTTATGCCAGGATTTACTAGAGAGTGTCA  
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTGTGGTTGGAGGAGAGAACCTTGTGGGGCTGCGTTCTCTTAGCA  
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACAGAGAAAGGAAAGGTCCTTCTGCTGTTGGCTGCACATCAGGAA  
GGCTGTGATGGGAATGAAGGTGAAACTTGGAGATTTCATTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA  
AGTAGAGAAGCTGCTCTGTGTGGTGGTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC  
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCAGGGAAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC  
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGTTTCGCCGGGGGCTGCTTGCCTGGATTTCGCCGGTGGTG  
GTTTTGCTGGTGTCTCTGTGTGCTATCTGTCTGTACATGTTGGCTGCACCCCAAGGTGACGAGGAGCAG  
CTGGCACTGCCAGGGCCAAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCCTCCTTCAGGAGTGGGAGGAGCAGCAC  
CGCACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGCAGCTCAGG  
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC  
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG  
CCTTTGATAGCTTACTCTACAGAGGTGTACCAGCTGGAGCTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTG  
AGGAAGGACAAAGCGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCTGCAGAGAACAGC  
CCCAATCACCGTCTTACACGGCCTCTGATTTATAGAAGGATCTACCGAACAGAAAGGACAAAGGACATTGTAT  
GAGCTCACCTTCAAAGGGGACACAAACAGCAATTCAAACGGCTCATCTTATTTGACCATTCAGCCCCATCATGAAA  
GTGAAAATGAAAGCTCAACATGGCCACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGCTGGACAAGTTT  
CGGCAGTTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAATGAAGTCAAAGGAATCTTGAAACACTTCCAAAGCTGCCAATTCAGGAACCTTACCTTCATC  
CAGCTGAATGGAGAAATTTCTCGGGGAAAGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCTCTC  
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG  
GTATTTTATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA  
CAGCAGCTGGTCATAAAGAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCGGTGAGAG  
TTCACTAATATAGGTGGGTTTGTCTGGACATCAAAGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC  
CACAGCAACCTCATAGTGGTACGGACGCTGTGCGAGGACTCTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG  
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG  
GTGTTCAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAACTCCGAGA  
GAAGGATTTGGGAGACACTTTTCTTTCTTTTGCATTTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAA  
GGACGCAAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTGTTGGCTTTTACAAAGAGA  
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG  
AGATTATAAGCCTAATGGTGTGGAGGTTTGTGATGGTGTTTACAATACACTGAGACCTGTTGTTTGTGTGCTCATTGA  
AATATTCATGATTTAAGAGCAGTTTGTAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAATATCAGAAAGGCAGGAGAGGAGATAGGCTTATTATGACT  
AGTGAGTACATTAAGTAAATAAAATGGACCAGAAAAGAAAAGAACCATAAATATCGTGTATATTTTCCCAAGAT  
TAACCAAAAATAATCTGCTTATCTTTTGGTGTCTTTTAACTGTCTCCGTTTCTTTCTTTTATTTAAATGCACT  
TTTTTCTCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC  
ATTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAAATGATGCCATAT  
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCACTGAGACATAGGGAAGGAATGGTTGTACT  
AATACAGACGTACAGATACTTCTCTGAAGAGTATTTTGAAGAGGAGCACTGAACACTGGAGGAAAAGAAAATGAC  
ACTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACTTTT  
CTCCTCAGAAGTAGGGACCGCTTTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTT  
AAAACAGGGTGTCTCCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTGT  
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTGCTACATGTTATCCACCCAGGCCAGGTGGAAG  
TAAGTGAATATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT  
CAAACTATTTTTTAAAAATAAATACAGTTAACAATAGAGTGGTTTTCTTATTATGTGAAAATTATTAGCCAGCACCAG  
ATGCATGAGCTAATATCTCTTTGAGTCTTGTCTGTTTGTCTCACAGTAACTCATTGTTTAAAGCTTCAAGAAC  
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG  
CCATGAATGGAAGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

## **FIGURE 72**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW  
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGDRSPPEKTQADLLAFLHSQVDK  
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPA  
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEKLNMAN  
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF  
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFPVLF  
SQQY NPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR  
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPQYKMCQSKAMNEASHGQLGMLVFRHEIEAHL  
RKQKQKTSSKKT

### **Important features:**

#### **Signal peptide:**

amino acids 1-27

#### **N-glycosylation sites.**

amino acids 315-319, 324-328

#### **N-myristoylation sites.**

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

#### **Amidation site.**

amino acids 377-381



## FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATCCA  
GAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA  
GAAGCAGAGATAAATACTCAGGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTCACTCCTC  
CCTCCCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATTTCCAGTCCCCTGCACCCCTTC  
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGGG  
GTCAACACTGGACGTATGAGGGCCCATGGTCAGGACCATTGGCCAGCCTCTTACCCTGAGTGT  
GGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTTGACCCTGATTTGCC  
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCAGCCTTTGGACCTGCACAACAATGGCC  
ACACAGTGAACCTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTTCCCCGAAAAATGTAGCT  
GCCAGCTCCACCTGCACCTGGGGTCAGAAAGGATCCCCAGGGGGGTGAGAACACCAGATCAACAG  
TGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCTGATTCCTATGACAGCTTGAGTG  
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCTGGGCATCCTAATTGAGGTGGGTGAGACTAAG  
AATATAGCTTATGAACACATTCTGAGTCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTC  
AGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAACAGCTGGGGCAGTACTTCCGCTACAATG  
GCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAG  
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAA  
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCA  
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGTTGGC  
TGTCTCTGCCTTCTCCTGGCTGTTTATTTTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAA  
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATTAAATTCTTCTCAGATAC  
CATGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGG  
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCCTTCCCCTGGACATCTCTTAGAGAGGAAT  
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAA  
ATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGGGAAGTTTGGG  
ATATACCCCAAAGTCCTCTACCCCTCACTTTTATGGCCCTTTCCCTAGATATACTGCGGGATCT  
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATATATTTGGAATTAAG  
TTTCTGACTTT

## **FIGURE 74**

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQ  
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT  
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP  
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLGTLFSTEEEPSKLLV  
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS  
VVFTSAQATTEA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

## FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTTGGGGACGGGCAGTTCCCTGTGTC  
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA  
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA  
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC  
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA  
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC  
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC  
AGGGCCCCCTCGCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT  
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT  
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT  
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA  
ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA  
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA  
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAATTTTTTGTGACTCTGAAG  
AAAACACGGAAGGTACTTCTCTACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA  
ACAGTCATTGAATATGAATATGATGTGAGAACCTGACATTTGTGCGGGCCCTGAAGAGCAGGA  
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG  
TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG  
CAGGAGCACACAGACTCGGAGGAGGGGCCGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA  
TCCCCAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG  
AGCCTTCTGAGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT  
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA  
TGTGCAGATGGAAAATGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAACAAGTGAG  
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT  
CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA  
GGTCTCTTAACAATGATGGTGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA  
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

## **FIGURE 76**

MSYNGLHQRVFKELKLLTLCSSSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY  
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK  
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFDRFFVPAEK  
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC  
DSEENTEGTSLTQQESLSRTIPDPKTVIEYDYDVRTTDICAGPEEQELSLQEEVSTQGTLLSQ  
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS  
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYQMEN

### **Important features:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 140-163

#### **N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

## FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC  
CTGGGAAGATGGCCCGGCCGTGGACCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC  
CAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC  
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC  
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC  
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA  
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAAGACCA  
TCGTGGAGTTCCACATGACGACTGAGGCCAAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC  
CCCCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTCTGTGA  
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC  
CCAATCTAGTGAAAAACAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC  
CTCCTGCAGCTGGTGAAGGTGCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTGA  
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA  
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCCTGGACAACATCCCG  
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
ATTCATGGTCCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC  
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC  
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGT  
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT  
TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG  
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA  
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG  
CCTTGGGATTTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC  
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG  
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAAGCTTG  
CCTGTGAAAAA

## **FIGURE 78**

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVKEKLTQELKDHNATSILQQPLLSAMREK  
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE  
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL  
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT  
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN  
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT  
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSQVPSLVKALG  
FEAAESSLTKDALVLT PASLWKPPSPVSQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

#### **Glycosaminoglycan attachment site.**

amino acids 412-415

#### **LBP / BPI / CETP family proteins.**

amino acids 407-457

## FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC  
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT  
CTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT  
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGCGTGCCAGCATTGTGACAGCAGTTGGCTTCT  
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC  
ACCTTCTGGGCCTGCCCCGTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT  
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC  
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTC  
ATTCTGTGTCCTGGAATCTTCATGGGATCTACGGGACTTCTACTCACCAGTGGTGCTGACAG  
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG  
CTGGAATCATCTCTGCTTTTCTGCTCATCCCAGAGAAATCGCTCCAACTACTACGATGCCTAC  
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTCAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG  
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT  
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG  
GCTAGTGTAACAGCATGCAGGTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC  
TCACCTTGCTGCTCCCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCCATTCCTTAAGCCA  
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA  
TCCCACTGACTGACCCTCTGTGATCAAAGACCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT  
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC  
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA  
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG  
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

## **FIGURE 8o**

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD  
IYSTLLGLPADIQAAQAMMTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL  
LGFIPVAVNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNY  
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-24

#### **Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

#### **N-glycosylation site.**

amino acids 190-193

#### **PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59



## FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT  
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCTGCCCCCTTTTCT  
GCCACCGCTGCTTCCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG  
GTTGATCTGTGGCCCCTGTGCCCTCCGTGTCCTTTTCGTCTCCCTTCTCCCGACTCCGCTCCCGG  
ACCAGCGGCCTGACCCTGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCTCTCTTGTCTGGGA  
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT  
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT  
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCT  
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC  
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG  
AGATCTTCAGTGCCCATGAGCTGTTCCCCCTCCCGCCTGCCAACCAGTGTGTCTCTGCAGCTGC  
ACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGAACCAGGCTGCCAGCACCCCTCCC  
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA  
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG  
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCCTCTGAGCTTCATCCCTCGCCACTT  
CAGACCCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT  
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC  
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC  
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG  
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC  
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC  
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG  
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG  
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCCT  
GGAACGTCTTCTAGCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG  
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA  
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 82**

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE  
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF  
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR  
HPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGKTY  
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVAGKCKICPEDKADPGHSE  
ISSTRCPKAPGRVLVHTSVSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH  
SQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPGAEGHGQSRQSDQDITKT

**Signal peptide:**

amino acids 1-25

## FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCAGCGAGAGCCTCTCC  
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTAATCCATCCGTCACCTCTCCTGTCA  
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC  
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG  
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG  
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC  
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT  
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT  
 ACAGGTGTCAGCACTGGGCTCAGTTTCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC  
 TCTGTCACTCCTCGGGCTGGTTCCCCCGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTCC  
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA  
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGTACAGATAG  
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA  
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAATCCAGGCGGAAGTGGACTG  
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCGAAACACGCAGTGAGGTGACTCTGGATCCAG  
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTG  
 CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA  
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA  
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT  
 ACATTAAATCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA  
 TGAGTGTGGGACCATCTCCTTCTCAACATAAATGACCAGTCCCTTATTTATACCTGACATGTGGTTTG  
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAAGTCCCATAGTCTCTGC  
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG  
 TGAGTCCTCCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT  
 TCTTCTTTAGGGATATTAAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA  
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG  
 CTGACATTACATTTAGTTTGTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG  
 AACCGTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG  
 ATCTTATTGATGACAGAGTGATCCTAATGGTTTGTTCATTATATTACACTTTTCAAGTAAAAAAA

## **FIGURE 84**

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFCFLSPKTNAEAMEVRFFRGQFSSVVH  
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ  
VSALGSVPLISITGYVDRDIQLLCQSSGWFRPTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL  
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK  
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF  
TRKSVVASQSFAQGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT  
LNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT  
PIVICPVTQESEKEASWQRASAIPESTSSESSQATTPFLPRGEM

### **Signal peptide:**

amino acids 1-17

### **Transmembrane domain:**

amino acids 239-255

## FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT  
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA  
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT  
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC  
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG  
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA  
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA  
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCAGGCACCCTGGAGTCCGGCTGCCCCC  
AGAATCTGACCTGCTCTGTGCCCTGGGCCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA  
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCCCTCATCCCACA  
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA  
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA  
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCT  
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA  
GAGGCTGACCTGTGCCCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTTCTCGGCTCTCAGCAGGTCTACCT  
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG  
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA  
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC  
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG  
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG  
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG  
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA  
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG  
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC  
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

## **FIGURE 86**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYPGPVVHGYWFREG  
ANTDQDAPVATNNPARAVWEETRDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY  
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPACEQGTTPMISWIGTSVSPLDPSTTRS  
SVLTLIPQPQDHGTS LTCQVTFPGASVTTNKT VHLNVSYPQNLMTVFQGDGTVSTVLGNGSSL  
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP  
LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE  
DANAVRGSSASQGPLETPWAEDSPDPQPPASARSSVGEGELQYASLSFQMKPWDSRGQEATDTE  
YSEIKIHR

### **Signal peptide:**

amino acids 1-15

### **Transmembrane domain:**

amino acids 351-370

## FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC  
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC  
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT  
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTTGA  
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG  
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG  
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC  
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT  
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG  
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT  
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC  
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT  
GGCCAGCGGGAATTCACTGCGGGATTGTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA  
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG  
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG  
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCTCCCAACCATGAGATCCCAAGGATGGAGAA  
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA  
AAAAAA

## **FIGURE 88**

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI  
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT  
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG  
EGKCWTDNGPVIPVVYDFGDAQKTASYYSYPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN  
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

### **Important features:**

#### **Signal peptide:**

amino acids 1-16

#### **N-glycosylation site.**

amino acids 163-167

#### **Glycosaminoglycan attachment sites.**

amino acids 74-78, 289-293

#### **N-myristoylation sites.**

amino acids 76-82, 115-121, 124-130, 253-259, 292-298



## **FIGURE 89**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC  
CGCCGCCCTTGTCCTCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG  
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG  
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT  
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC  
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT  
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGGCCCTTCCAGCTG  
TCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTCGATACCTTCA  
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG  
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG  
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA  
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

## **FIGURE 90**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACPLTFTTPEEYDKQDIQLVAALSVTLGLFA  
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV  
TVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

10063732

## FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCCAACTGAC  
TCGCTGCTGCTTCGTGTTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG  
GTCTCTCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT  
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCAATCTCCTAGGGCTGCT  
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCAAGCCCCC  
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG  
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT  
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG  
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG  
GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG  
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG  
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC  
CATAGTGATACCCCCCTACTACCCATCTGGGTGACCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC  
TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC  
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG  
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG  
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC  
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACA  
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG  
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGG  
CATGAGGCTGAAGTGGAACCCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC  
CAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG  
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC  
CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG  
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTTCTCCCGGATCTGGATGGCGC  
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGAGCATGTGCTGGATCTGTTC  
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTGTGAAACCGCTGATTGCTGACTTT  
TGTGTGAAGAATCGTGTTCTTGAGCAGGAAATAAGCTTGCCCCGGGGCA

## **FIGURE 92**

MQLTRCCFVFLVQGSLLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL  
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF  
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEEKVERGRRTSLCTHDP  
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

## FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC  
CTTTATGTCTTCACCATCGCCATCGAGCCGTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT  
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA  
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA  
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC  
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA  
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT  
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT  
ATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC  
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG  
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG  
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGTTTCTTCTTTACAACCAGCGCTCCAGAT  
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT  
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

## **FIGURE 94**

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG  
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV  
FSFVNTLSDSLPGPTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT  
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

**Glycosaminoglycan attachment site.**

amino acids 120-123

**Sodium:neurotransmitter symporter family protein**

amino acids 31-65

2025-03-27 15:00:00

## FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG  
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC  
TCGGTCATTACCACAGCTCAAACCTGCCTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG  
GAACTACTACCAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAAACA  
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC  
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC  
CAATTTTGTGACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC  
TTCACGAGCCTCATCATCCATTCCCTTGTTCCTGGGAGGCATCCTGCCCCACAGTCAGGCAGGGGC  
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG  
GAACCCAGCAGGCCGCCTCCCAACTCCCACTGGCACAGATGACGACTTTGCAGTGACCACCCCT  
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA  
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC  
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT  
TACCTGAAAATATTCTTGAAATTTGAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA  
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT  
TAAACATATTTGAAAACCTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 96**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLISLIPLTQM  
LTLGPDHLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPFVFTQLGAQGTILSSEE  
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLETPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16



## FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT  
 CTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTGACC**ATG**  
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGGCCAGCCTGC  
 AGAGCTGTCTGTGGAAGTTCAGAAAACATATGGTGGAATTTCCCTTTATACCTGACCAAGTTGC  
 CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAAGGCAACTGAG  
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGTGCTGGTGACCAGGGCCCTGGACCGAGAGGA  
 GCAGGCAGAGTACCAGCTACAGGTCACCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC  
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC  
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCCTGGCATCCCCCTTCTCTTCTTGGGCTTCAGACCG  
 GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC  
 CTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC  
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA  
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGT  
 CCTTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTA  
 CACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAA  
 TGCAAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGTCTCC  
 AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG  
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT  
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA  
 ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC  
 CAGGTGGACCCCACTTCAGGCAGTGTGACCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT  
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCAGTGTGAAG  
 TCGAAGTCGCAGTCAAGATATCAATGATCACGCCCCCTGAGTTTCACTCTCCAGATTGGGCCCT  
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGCCATGCTAACAGCCATTGATGCTGA  
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG  
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG  
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC  
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGTGCCACCCCCCAAGTTGG  
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTCTCTGCTGACCATC  
 CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT  
 CTGCATTGAGAAATTCTCCGGGGAGGTGCACACCGCCCACTCCCTGCAGGGCGCCAGCCTGGGG  
 ACACCTACACGGTGCTTGTGGAGGGCCAGGATACAGCCCTGACTCTTGCCCCCTGTGCCCTCCCAA  
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT  
 GGCCAGTGGGCACGGTCCCTACAGCTTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC  
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGT  
 GAACACATAATCCCCGTGGTGGTCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT  
 CGTGTGTGCGTCAACGTGGAGGGGCAGTGCATGCGCAAGGTGGGCCCATGAAGGGCATGCCCA  
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC  
 ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT  
 GAAGGCGACTGT**TGA**ATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCTCTGGCTCCATCTGAG  
 TCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA  
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGCATGTCTGCAGAGCCTGGACACCACTT  
 TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGTTGCCCAATAATAAGCCCCAGAGAA  
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

## **FIGURE 98**

MVPAWLWLLCVSPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT  
EGPFAMDPDSGFLLVTRALDREEQAQYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI  
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG  
STSLDHALERTYQLLVQVKMDQASGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ  
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL  
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA  
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG  
PISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY  
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT  
IQPSDPISRTLRFSLVNDSEGLWCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS  
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP  
REHIIPVVVSHNAQMWQLLVRVIVCRNVEGQCMRKVGRMKGMPTKLSAVGILVGTTLVAIGIFLI  
LIFTHWTMSRKKDPDQPADSVPLKATV

### **Signal peptide:**

amino acids 1-18

### **Transmembrane domain:**

amino acids 762-784

## FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG  
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA  
TCTCTTACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC  
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG  
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAATTTCCAATGAGACTAGCACC  
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG  
TGTGACCTCCAGTGGGGTCCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA  
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG  
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG  
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCACCAACT  
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC  
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC  
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT  
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA  
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG  
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT  
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTACCAATTCTGAGTCCAGCACACCCTCC  
AGTGGGGCCACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCACACAGCCAC  
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA  
CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA  
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG  
CACAGTGTCCAGTGGGATCAGCACAGTACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA  
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA  
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAGCCTGGTGGGTCCCTGGT  
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC  
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT  
CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCAGGCCCCAG  
GTGGAGTCTTAAGTGGTTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA  
ACAGCGGGCCCCTGAGCAGCCCCGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA  
CCCAAGACCTGGTTTCTTTTATTTCATCCAGGAGACCCCTCCAGCTTTGTTTGGATCCTGAA  
AATCTTGAAGAAGGTATTCTCACCTTTCTGCCTTTACCAGACACTGGAAAGAGAAATACTATAT  
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG  
CCCCGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC  
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 100**

MRMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS  
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTSSGASTATNSESSTPSS  
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA  
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES  
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS  
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA  
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES  
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA  
AVGLFAGLFFCVRNLSLRNTFNTAVYHHPGLNHGLGPGPGNGAPHRPRWSPNWFRRPVSSI  
AMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532

## FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA  
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGGGCGTGCCCTGCTTGTCACAGGTG  
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT  
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT  
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT  
GGTGTCATGGAAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT  
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC  
CGCCTTGTTGGCAACATGGCTCACCCCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT  
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA  
GAATGGACTGGATGTTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT  
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTTGGTTGGGCCATGCAGGCTGCTGCCTATAT  
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCTGAAGACATGATTGATTACTTTTGTG  
ATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG  
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG  
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG  
ATATCACTGTGGCGTATCCTCACAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT  
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA  
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAGAGGCTGCGTTCCTTCTATCAAG  
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTC  
CTTGTTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCAT  
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA  
GAATATTTGGTGGACTGGAGATCATAGAACTGCATGTTACCGACTTTTACACAAACAGCCACAT  
TTAAATTCAAAGAAAAATGAGTAAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG  
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACT  
GCCATCATTATTTGTTAAAGATATTTTGCACTTAATTTTGTGGGAAAAATATTGCTACAATTTTT  
TTTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT  
GGGCCAGAATATATTAAACAATCATCAGGCTTTTAAA

## **FIGURE 102**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIFMLSP  
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM  
FLWNCMLRYSYLRLKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP  
LQLLIFPEGTDLTENSKSRNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDI TVA  
YPHNIPQSEKHLQGDFFPREIH FHVHRYPIDTLPTS KEDLQLWCHKRWEEKEERLRSFYQGEKNF  
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG  
LEIIE LACYRL LHKQPHLNSKKNE

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 44-63, 90-108, 354-377

## FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC  
TCCAAATCATCCATCCACCCCTGCTGTCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA  
TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT  
CACTGGACCGGGCAAGTTTGTCCAGGCCCTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT  
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC  
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA  
GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG  
ACATCGGCCTGTATGGGTGCTGGTTTCACTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG  
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA  
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC  
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC  
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA  
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA  
TTTTACTCGGGTACTCTGTGGTGGCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC  
AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAAATTGAG  
AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG  
TTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA  
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA  
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGGGATGACGTAGACAGGGGGAAGAACA  
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT  
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT  
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC  
TGACATGTCAGTTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAAG  
GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATGACAGACAGAGAAGACCCTGCTTAAAGGGC  
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT  
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCCTCTCTTTAGGGAGCTGAGGTTCTCTGCCC  
TGAGCCCTGCAGCAGCGGCAGTCCAGCTCCAGATGAGGGGGGATTGGCCTGACCTGTGGGAG  
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA  
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCAGGCTCCTCATTGCTAGTCACGG  
ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT  
TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT  
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC  
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAGCAT  
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC  
CTCATTAACACAGACACAAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAATAATTTA  
AAGATGATATATAACTACTCAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAAATTTAAAT  
ATCAACCAGTGTAATTGAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

## **FIGURE 104**

MAFVLILVLSFYELVSGWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH  
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR  
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI  
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK  
SKGKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF  
TRKSVVASQGFQAGRHYEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT  
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG  
TPIFICPVSWG

### **Signal peptide:**

amino acids 1-17

### **Transmembrane domains:**

amino acids 131-150, 235-259



## FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG  
TTTGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA  
CTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC  
TGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTACAGAAATGAGCCAGAGACTTGAAT  
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC  
AGTCAACAGAAGCATGGAGTGTGGGTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGA  
AACTGTAGATAAAATTTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG  
ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA  
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG  
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC  
TTGTGAGTGTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA  
ATAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA  
TGACTATGATATTTCTCTGTCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT  
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT  
GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC  
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT  
GCCAGGGTGAATCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG  
AGCTGGGGAGATGAATGTGCGAAACCCCAAGCCTGGTGTATATACTAGAGTTACGGCCTTGGGGACTG  
GATTACTTCAAAAACCTGGTATCTTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTGTTTTTTG  
GGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA  
ATAAACTGTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA  
GATCAACTCTGTCATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC  
ATTACAGCCTGTATTCAATTTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT  
GCATATATACAATTTGAAGCACTCCTTTTCTCAGTTCCCTCAGCTCCTCTCATTTCAGCAAATATCCATTT  
TCAAGGTGCAGAACAAAGGAGTGAAGAAAAATATAAGAAGAAAAAATCCCTACATTTTATTGGCACAGAA  
AAGTATTAGGTGTTTTTCTTAGTGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAGACA  
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAATAAGTTAAGGAAGTCCAGAAAGAAGCCAAG  
ATATATCCTTATTTTCATTTCCAAACAATACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT  
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG  
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

[illegible]

Transmembrane domain:

APP ID=10063732

## FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTTG  
CCCCGATGAGCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC  
GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT  
ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA  
TCTGCTCTTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTGGATGAGGGACTACCTAA  
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCTT  
GTGGTTCATCCGGTCATGATTGCTGTTTGTCTGTTTCTTATCATTGTGGGGATGTAGGATATTG  
TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGTCTTGTCATTTTCT  
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAAGTTATGGTTCAGTACAATGGTCA  
GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA  
TGCTTGGAATTTTTTTCAGAGAGAGTTTTAAGTGTCTGTGGAGTAGTATATTTCACTGACTGGTTGG  
AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTAGAGAATTCCAGGATGTTCCAAA  
CAGGCCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT  
TTTGAGAGGAACCAACAACATGCAGGTGCTGAGGTTTTCTGGGAATCTCCATTGGGGTGACACAAA  
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG  
ACAGACCAAATGATGTCCTTGAAGAATGACAACTCTCAGCACCTGTCATGTCCCTCAGTAGAACT  
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT  
TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACT  
TGTAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA  
TAACACCTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTCATGTCATAAGTC  
ACCACCTGGACAATAATTGATGCCCTTAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC  
TGTTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA  
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGTTGATTTACTTCTACCAA  
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTTATTA  
CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCAATATTGGTGACT  
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACAT  
TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT  
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT  
ATTTTTAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGTCCTGGGCTTATATTACACATATAAC  
TGTTATTTAAATACTTAACCCTAATTTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC  
AGAATGTAGTCTGGTCTTTAGGAAGTATTATAAGAAAATTTGCACATAACTTAGTTGATTTCAGA  
AAGGACTTGTATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA  
GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCACAGAAAA  
TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAAACATGTGACAATTTAGAGATT  
CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTACAA  
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTAT  
TTCTCAGAATATGGAAGAAAATTAATAATGTGTCAATAAATATTTCTAGAGAGTAA

## **FIGURE 108**

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP  
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMT  
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ  
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM  
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

### **Signal peptide:**

amino acids 1-33

### **Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248

## FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCTCTGATAAAGCCCTACCAGTGCT  
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAGTGCCTTGAAAGAGAAGGGGACAAAGGAACA  
CCAGTATTAAGAGGATTTTCCAGTGTTCCTGGCAGTTGGTCCAGAAGGATGCCCTCCATTCTGCTTCTCACCTG  
CCTCTTCATCACAGGCACCTCCGTGTCAACCGTGGCCCTAGATCCTGTTCTGCTTACATCAGCCTGAATGAGC  
CCTGGAGGAACACTGACCACCAAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG  
TGGTACCACCTTACGGGCATGGCGGGAGATGCCATGCCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA  
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCAGCGCATTGTGCAACGCCAGGCTTGTGCCAGCT  
TCAATGGGAAGTGTGTCTCTGGAACACCACGGTGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT  
CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG  
CAGCTGCTCAGATACCAGCGAGTGCACATGCCGTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG  
ATGAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT  
GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA  
TGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCGGGGCTGGTGTGT  
CTGAGGATAACCACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG  
CTGGTTGGTGGCTGGAGCTCTTCTGACCAACACCTCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT  
CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA  
CAGGTCTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG  
ACCTGCGAGTTTCCACGCCGTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT  
CATGAGCCGAAATCATGGGATCTTCCCATTTACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC  
GGGAAGCTCTGCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC  
TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCCTGAAATACTACCTCAT  
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG  
TCCCTGTCTTCAAGTTTGTGGGCAAGACCACAAGGAAGTGTTCGCACTGCCGGGTTCTTGTCTGTGGAGTG  
TTGGACGAGCGTTCCCGCTGTGCCCAGGTTGCCACCGCGAATGCGTCGTGGGCAGGAGGAGGACTCAGC  
CGGTCTACAGGGCCAGACGCTAACAGCGGCCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC  
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC  
TGGGTTGAGACTTCACACTGTGAGTTGAGACTCCAGCACTCACTCTGATTCTGGTCCATTGAGTGGGCA  
CAGGTACAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTTCTAGGGTTGAAAACCTAACTGTCCA  
CCCAGAAAGCACTCACCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGAATCAGAC  
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA  
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA  
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

## **FIGURE 110**

MPFFLLLTCLFITGTSVSPVALDPCSAYISLNEFWRNTHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP  
TFCIPENHCGTHAPVWLNNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV  
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV  
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSHDNHTCQVPVLCKSNAIEVNIIPRELVGG  
LELFLTNTSCRGVSNNGTHVNIILFSLKTCGTVDVVDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT  
CEFPRLYITISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLKRLDSLYFGIEPVVHV  
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDKHEVFLHCRV  
LVCGVLDERSRCAQGGCHRRMRRGAGGEDSAGLQGQTLTGGPPIRIDWED

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-16

#### **N-glycosylation sites.**

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

#### **Tyrosine kinase phosphorylation sites.**

amino acids 411-418, 443-451

#### **N-myristoylation sites.**

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
522-528, 531-537

#### **Aspartic acid and asparagine hydroxylation site.**

amino acids 197-209

#### **ZP domain proteins.**

amino acids 431-457

#### **Calcium-binding EGF-like proteins.**

amino acids 191-212, 232-253

## **FIGURE 111**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG  
 GCCTCCTCCAGCCAGTGTCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT  
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT  
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC  
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT  
 GGTGTGCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA  
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC  
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC  
 AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG  
 GCTACAGCAGAGCTGTGGAGATTGGCCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG  
 GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTG  
 TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGG  
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA  
 GCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG  
 CTTCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAGACAATGACA  
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT  
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA  
 GATGTCTGACATACTGTGTCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGT  
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT  
 GACAGTGGTGGGCCCCGTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG  
 CTGCGGGGGCCCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT  
 GGAAGGCTGAGCTGTAATGTCTGTGCCCTTTGCAGTGTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT  
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT  
 TTCTTGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGGCGCCCAGAGGAAGTCA  
 GCAGCCCTAGCTCGGCCACACTTGGTGTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT  
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAGCC  
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA  
 GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT  
 ACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA  
 AAAA

## **FIGURE 112**

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG  
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN  
FTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL  
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNNKVRAGSDKL  
GSFPSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG  
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQDSSGGPLMYQS  
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

**Transmembrane domain:**

amino acids 32-53 (typeII)

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCTCAGCCTCCCAAGGTGCTGTGATTA  
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT  
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC  
TGTTTTTTGTTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT  
GGCCCAAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCCTCGAAG  
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT  
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC  
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC  
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA  
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCCTGCACATCACCTG  
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC  
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG  
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTGTGCCAAATTATGGGTGAGAAAAGATG  
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTCAGTGGACATGTGGGGGAAGGGCTG  
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT  
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA  
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTCTCCTCCC  
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG  
GGCTAATGGCTCAGTGTTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT  
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG  
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAGGAGCTAGAGCTTGTTT  
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA  
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT  
TTTGGGAGGCCGAGGCGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG  
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC  
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA  
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA  
TGTTTATTTGTAA

## **FIGURE 114**

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC  
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

**Signal peptide:**  
amino acids 1-15

10063732

## FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA  
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA  
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCCT  
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT  
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA  
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG  
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT  
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG  
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA  
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG  
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT  
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA  
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTGATCCCATGCTGGATGAGAGAG  
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA  
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG  
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTTGAGCTCAAATGCTTAAAC  
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT  
GGTAGCCAGCTCTCCAGAATTACTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA  
TTATCACCAAAAAAAAAAAAAAAAAA

## **FIGURE 116**

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME  
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIFYVGLQKCFIKTQIKVIP  
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE  
LQDFEEEGEDLHFPANEKKGIEQNEQWVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD  
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-40

#### **Transmembrane domain:**

amino acids 25-47 (type II)

#### **N-glycosylation sites.**

amino acids 94-97, 180-183

#### **Glycosaminoglycan attachment sites.**

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

#### **N-myristoylation sites.**

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

#### **Microbodies C-terminal targeting signal.**

amino acids 315-317

#### **Cytochrome c family heme-binding site signature.**

amino acids 9-14

## FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA  
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCCTGTCCATCCTGGGGCT  
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT  
CCGTGTTCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC  
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCTT  
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG  
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG  
TCTGTGTTTGGCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG  
GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGCTCTGTTTCGTGGGCTGGGTGCGTGGAGGCC  
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCAACTACAAA  
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT  
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC  
CTTCCAAGCAGCACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA  
CCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGTCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT  
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA  
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT  
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC  
TTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCAAGAAAACCTTTTGAAGGAAA  
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGCTAGTAATAA  
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT  
ACACTGTGATCTTAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTCTACTGCTG  
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTCTGT  
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCTAAATATAGTTAA  
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG  
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC  
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT  
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA  
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT  
TGGGGCTGCAGTGAGCCATGATCACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA  
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATCTTTAA

## **FIGURE 118**

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP  
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL  
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA  
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY  
V

### **Signal peptide:**

amino acids 1-23

### **Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

## FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGGAGTCC  
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG  
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT  
GAAAACAACATCGTGGTTTTTGAAAACCTCTGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA  
CATCAGGATGCAGTGC AAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG  
GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC  
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT  
CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTCTATA  
ACTCAATAGTGAATGTTGCCCAAAACGTGAGCTTGAGAAGCTCTCTACTTAGGATGGACCACGGCA  
CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA  
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG  
TCTACTCCAGAAGTCAGTATGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG  
ACAAAAATCTATATTACTTTCTCAAAATGGACCCAAAGAACTTTGATTACTGTTCTTAAGTGCCT  
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGAGAATGAGATA  
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA  
CTCTTTTTATCATTACTTCAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTTCTCC  
ACGACATAGCATTATGTACATAGATGAGTGTAACTTTATATCTCACATAGAGACATGCTTATATGGT  
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA  
ATCATGGATAGGGTTGAAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA  
TTTATAATGAAGATTTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATGAATGGCTTTCTGATAT  
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAAGTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT  
CTTGTGTATTAAATTAACATTTTAAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAAGTGCCT  
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGATGGTTTTAGGAAAGTG  
AAAATATTTTTGTTTTGTATTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT  
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA  
TGTCTTGGTTTTCAATTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTGAGAACTTCACCT  
GCTCCTATGTGGGTACCTGAGTCAAAATGTCATTTTGTCTGTGAAAAATAAATTCCTTCTTGTA  
CCATTTCTGTTTAGTTTTACTAAATCTGTAAATCTGTATTTTCTGTTTATTCCAAATTTGATGAA  
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT  
TATACATTTATATTAATAAATGTACATTTTCTAATT

## **FIGURE 120**

MATHALEIAGLFLGGVGMVGTVAVTVMQPWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK  
IYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG  
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY  
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

### **Signal peptide:**

amino acids 1-17

### **Transmembrane domains:**

amino acids 82-101, 118-145, 164-188



## FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG  
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGCAGCC  
GGGAGCCATGCGACCCCAGGGCCCCGCCCTCCCCGAGCGGCTCCGCGGCCTCCTGCTGCTCC  
TGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG  
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC  
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTGCGGGATG  
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACCTAC  
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT  
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA  
GAAATGCATGCTGTGAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC  
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG  
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGAATGTTGCTATCTGGG  
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC  
ATTATTGAAGAACTACCAAAATTAAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTTATTATGCC  
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG  
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG  
CTTCAATCAAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT  
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTAA  
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT  
TAAATAAAAATTATTTCCAACA

## **FIGURE 122**

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKKGKQKAQLRQREVVDLYNGMCLQGPGVPGR  
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK  
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS  
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

### **Signal peptide:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 195-217

## FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA  
CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA  
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA  
ACCGCCCCCTCCCCACCCCCCAAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC  
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG  
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT  
TCTTCTGGGTGTGAGCGAGCCCTGACTACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG  
GCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTTCAATG  
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCCACTGTCTTACTGACAATG  
CTTCTTCTGCGGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA  
ATCTCAGAAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTTCCCTTCGCT  
ATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC  
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGA  
GCTGATTCTTAGTTCCAATAGAATCTCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT  
TACGGAACCTTGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG  
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA  
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG  
TCTTTGCTGGCATGATCAGACTCAAAGAACCTTCACCTGGAGCACAAATCAATTTTCCAAGCTCAAC  
CTGGCCCTTTTTCCAAGGTTGGTCAAGCTTCAGAACCTTTACTTGCACTGGAATAAAATCAGTGT  
CATAGGACAGACCATGTCTTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA  
TCGAAGCTTTTCACTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT  
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAG  
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAACCTGGCTGAAAAGTT  
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG  
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG  
GGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT  
TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT  
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGGTTAT  
CTACGTGTCTGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA  
GGCAGAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA  
GATTATAAACCACCAACACGAGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA  
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAGCT  
GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC  
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCATTTCATAATACT  
GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT  
GAACTCCGGTTAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT  
GTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

## **FIGURE 124**

MGENVIRLLSGSAVALVIAPTPLLTLMLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG  
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR  
PVTNLRNLDSLQNQLHSLGSEQFRGLRKLSSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS  
LARNVFAGMIRLKEHLHLEHNQFSKLNLALEPRLVSLQNLQWKNISVIGQTMSWTWSSLQRLDL  
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN  
WLKSFKGLRENTIICASPKELOGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE  
SKPPLPPTVGATEPGPETDADAHEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR  
SLMRRHRKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-33

#### **Transmembrane domain:**

amino acids 420-442

#### **N-glycosylation sites.**

amino acids 126-129, 357-360, 496-499, 504-507

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 465-468

#### **Tyrosine kinase phosphorylation site.**

amino acids 136-142

#### **N-myristoylation sites.**

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

## FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG  
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT  
AGCAACCTTTCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC  
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC  
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT  
GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG  
ATGAGCATTATCCCCTTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT  
TTAGCCAATCCAAGTACCTAGTGAAGGTTGAGATGCAAAATGGAAGGAAAAAGGAACTGGAAGG  
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC  
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA  
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC  
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA  
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG  
ACTGACTGCTTGATTGAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACC  
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA  
TGAGTGGAGTCAGTCCATTTTAA

## **FIGURE 126**

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES  
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVKSEDEHYPLWKS  
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP  
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP  
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTTPWSMVFWLTYEKIREMSGVSPF

### **Transmembrane domains:**

amino acids 25-38, 130-147, 233-248

## FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC  
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG  
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCCGTGGCC  
CGGGCCCTGGTCCAGCAGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA  
GCTGGCTGCTGAATGTAAGAGTGACAGGCTACCCCGGACTTTGATCCCCCTACAGATGTGACCTAT  
CAAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC  
TGCATCAACAATGCTGGCTTGGCCCGGCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA  
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA  
AGGAGCGGAATGTGGACGATGGGCACATCAATACATCAATAGCATGTCTGGCCACCGAGTGTTA  
CCCCGTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT  
GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG  
AGACACAATTCGCCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA  
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTATCTACGTCTCAGCACCCCCGCACA  
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC  
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT  
CACGGGATACCACTTCCGTGCCACACCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA  
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC  
TAATTGTTTTACTTGTTAACTGTTCTTGTCGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG  
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG  
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC  
AGACTTCCTCCTCTGCCTGCCCCACTGCACCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC  
AGCCCAGTCTTGGCTTCTTGTCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG  
CAGAACACCAGGGCCTGGCCCAGTGGATTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA  
AAAAAAAAA

## **FIGURE 128**

MARPGMERWRDRLALVTGASGGIGAARALVQQGLKVVGCVGNIEELAAECKSAGYPGTLI  
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLMSGSTSGWKDMFNVNVLALSICTR  
EAYQSMKERNVDDGHIININMSGHRVPLSVTHFYATKYAVTALTEGLRQELREAQTHIRATC  
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

**Important features of the protein:**

**Signal peptide:**

amino acids 1-17

**N-myristoylation sites.**

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,  
199-205

**Short-chain alcohol dehydrogenase.**

amino acids 30-42, 104-114



## FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGTCTTCTCCTCAGCCTCCTGCCGGTGGCCTACACCAT  
CATGTCCCTCCACCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC  
ACCTCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTCTGGTTTCATGC  
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA  
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG  
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA  
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA  
TATATAATGATCAAATCAGGGTAACCTGGGATATCCATCACATCAAACATTTATTTTTATTTCTTT  
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC  
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT  
GCACCACAATGCCCACTAATTTTTGTATTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG  
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTATGATTACA  
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTGGGAACTTTGAAATTAT  
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT  
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC  
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG  
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG  
TTCCATCCATGTTGCTGCAAATGACAGGATTTGTTCTTAATTTCAATTAAATAACACACATG  
GCAAAAA

## **FIGURE 130**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV  
KGGHTLGESPMFPKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

## FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC  
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG  
TCCATCTGTGTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG  
GAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCTTCAGAT  
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTCCTACCAACCT  
CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA  
AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA  
TTCCGAGACAGCAACTATCTCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTTT  
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG  
GTCTCACTAGTCTAAAACGCCTGGTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT  
TTCTTCAACCTAGTTAAATTTGACAGAGCTGTCCCTGGTGCAGAAATTCCTGACTGCTGCACCAGTAAACCT  
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT  
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT  
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG  
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA  
AGGTTCTGTTGGGATGGCTATTAAGGATCTCAATGCAGAACCTGTTTGATTGTAAGGACAGTGGGATTGTAAGC  
ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCCAAGGACAGTGGCCAGCTCCAGTGAC  
CAAAAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA  
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCAATATCTCTTGGAACTTGCTCTACCTATG  
ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTGGATCTATAACAGAAACAATTGT  
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCCATATAAGTATGCATGGTTC  
CCATGGAAACCAGCAACCTCTACCTATTTGATGAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT  
CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC  
TTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG  
TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAGAAAGGATGACTAT  
GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAATCAGGGAACTTCTTTTCAGATGTTACCAAT  
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATAATTCCTCCTAATGGAATGAATCTGT  
ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC  
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG  
GT

## **FIGURE 132**

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL  
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTTITYDSLSKIPYL  
EELHLDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL  
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAFPVNLPGTNLRKLYLQDNHIN  
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLNDITQLILRNNPWYCGCKMKWVRDWLQSLPV  
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQQGWPAFVTKQPD  
IKNPKLTKDQQTGSPSRKTTITTVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET  
IVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVC IETETAPLRMYNPPTTLNREQEKE  
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS  
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSSNRSYRDSGIPDSHSHS

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-28

#### **Transmembrane domain:**

amino acids 531-552

#### **N-glycosylation sites.**

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

#### **Tyrosine kinase phosphorylation site.**

amino acids 515-522

#### **N-myristoylation sites.**

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,  
640-645

#### **Amidation site.**

amino acids 567-570

#### **Leucine zipper pattern.**

amino acids 159-180

#### **Phospholipase A2 aspartic acid active site.**

amino acids 34-44

## FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGCA  
GCCCTGCCGGGCCACTTGTCTTATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG  
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG  
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT  
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT  
GGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT  
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCAGGAGCCCTGC  
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC  
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCAGCTGACCCTCCAGCCCGG  
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGTTGTACCCACGTTCCGGC  
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG  
GACAGCAGCGAGCCCTGCGGCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC  
AGGCTACTGCCTGTCCCACCACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG  
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC  
AGAGCTGAGGCCATCGGATACGCTACCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG  
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA  
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA  
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT  
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC  
AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT  
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT  
CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTTGGGATTATAGGTGTGAGCCACCGTGTCTG  
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC  
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG  
GTGGCTTCCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG  
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC  
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG  
ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG  
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA

## **FIGURE 134**

MSARGRWEGGRRACRGSGLGLARAQGAERVTSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD  
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV  
GMLGEKLEAAIQRSYLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE  
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQSQSD  
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCGF  
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILLP  
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPTSGSQSVGL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 39-56

#### **Tyrosine kinase phosphorylation sites.**

amino acids 149-156, 274-282

#### **N-myristoylation sites.**

amino acids 10-16, 20-26, 63-69, 208-214

#### **Amidation site.**

amino acids 10-14

#### **Glycoprotein hormones beta chain signature 1.**

amino acids 230-237

## FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCAATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG  
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG  
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGAGTTGTTCCAGGGGTGAAGC  
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG  
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC  
AGCTTACAGATTTGATCCCGTTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG  
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT  
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT  
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC  
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT  
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG  
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGTCAC  
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT  
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT  
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTGTCTGTAT  
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT  
ATGAACTACTATACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA  
AAAAAAAAAAAAAA

## **FIGURE 136**

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA  
RVLVDGEEHVGFLLKTDGSEFVVDIPSGSYVVEVVSAPYRFDVPRVDITSKGMRARYVNYIKTSE  
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME  
QSMNMLNSNHELPDVSEFMTRLFFSSKSSGKSSSGSSKTGKSGAGKRR

### **Important features of the protein:**

#### **Signal sequence:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 161-182

#### **N-glycosylation site.**

amino acids 184-187

#### **Glycosaminoglycan attachment sites.**

amino acids 37-40, 236-239

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 151-154

#### **N-myristoylation sites.**

amino acids 33-38, 36-41, 38-44, 229-234

#### **Amidation site.**

amino acids 238-241

#### **ATP/GTP-binding site motif A (P-loop).**

amino acids 229-236



## FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGCGATTTCTCCCCAGTTCCTGTGGGTCTGAGGGGA  
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA  
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCAGTGCAGTGCCTGCTGTTCCAGGC  
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAGGAGCCACGACCTGTGC  
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT  
TTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC  
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG  
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG  
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA  
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA  
AACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA  
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG  
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC  
TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG  
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA  
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG  
CTTCCTTGATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT  
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA  
ATGTATTTATTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTATAACCTG  
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAATAACCTTGTAATTCAGAAAGAGTGG  
CTAGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT  
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGTATGTGAATTGCAC  
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG  
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAAGTTTTCTTGCATACCAAAAAAAA  
AAAAA

## **FIGURE 138**

MRQFPKTSFDISPEMSFSIYSLQVPAVPGITCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF  
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG  
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS  
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET  
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGCCATCCAGCC  
TAGCGTGTCCACGATCGGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC  
TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCTCTGCCAGAGCGGAA  
CACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC  
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG  
GGTCAAAGGGTGTGAAATTTATGCCCTACACAATTACCTTGTGGAAAAAGGAGCATCTCACAGT  
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG  
CCTTCTCGGCTTGTGCGACGTCATCAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA  
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA  
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTCGTGTGATGATTACACAGAGGT  
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC  
TCCACTACCTGGGCTGGACCACATTGGCCACATTTAGGGCCCAACAGCCCCCTGATTGGGCAG  
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTGAAGGAGAGAGA  
GACGCTTTTACCAATTTGCTGGTTCTTTGTTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG  
GGGCTCTCTCCACCGAGGAGGTGAATACACCTCTGATTTAATCAGTCTGCGTTTGAAAGGAAA  
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC  
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCCTATCCAGTTGTGGAAGGAAGAC  
CAATGAGAGAGCAGTTGAGATTTTACATTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG  
AATGTGCCGTCTATGAAAAAGATCCTGGGTTTGGAGCAGTTTAAATGTGAGAAAGATTGCATGG  
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCTTATCAACCTGGGCTCCAAGG  
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCCTGAGTGCACAAGTGGCCCCAG  
TTCTCACCTGCTCTGCTCAGCGTCCCACAGGCACCTGCACAGAAAGGCTGAGCTGGAAGTCCCA  
CTGTCATCTCTGGGTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACAGT  
CATTGTGTGCACCTCAGCTGAAAGTTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT  
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC  
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCGACACAGGTG  
TTCACATCTGTGCTGTGAGGTGAGATGCCTCAGTTCTTGGAAAGCTAGGTTCTGCGACTGTTAC  
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA  
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA  
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACTCATCCTGCCACCCCAAGATGCATCCT  
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATTCTTAGTCTTGGCC  
TCGGACACCTTCATTCTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGCGGATGGTCAC  
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC  
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTGAGAGCCCTGTGCTGAGTGT  
CTGACCGAGACACTCACAGCTTTGTGTCATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG  
CCACGCTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG  
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

## **FIGURE 140**

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGHGAEPPEPSAGASSNWTTLPPPLE  
SKVVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF  
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWWKLFPKHFVEYDGTTSFFVSDYTEVDNNV  
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP  
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-34

#### **Transmembrane domain:**

amino acids 58-76

#### **N-glycosylation sites.**

amino acids 56-60, 194-198

#### **N-myristoylation sites.**

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,  
275-281, 278-284

#### **Amidation site.**

amino acids 154-158

#### **Cell attachment sequence.**

amino acids 205-208

## **FIGURE 141**

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA  
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC  
TCCCTATAGAAAACAACGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC  
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT  
AGCAGTTCAGATAAAAACACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT  
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC  
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT  
GGCTGCCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTATAGGGCTCAGGTGGGCTCCTGGA  
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT  
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA  
AGCTGAAATGAGCCCCAGTGAGGTGAGCGATTAGGAACTGCCCCATTGAACGCCTTCCTCGCTA  
ATTGAACTAATTGTATAAAAAACCAAACCTGCTCACT

**FIGURE 142**

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI  
FFALASSLSSASAEGKSPILLGVSKGEFLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI  
FYRAQVGSWNMLESAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

## FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA  
TCCAGGATCCTGTCCTTCCCTGTCCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG  
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTGTGGG  
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG  
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC  
CTGGCATTTCACGAGCAAAGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA  
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG  
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT  
GGGGAGAACTAGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG  
AGCTGAACAATACTTTACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC  
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT  
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT  
TTGGACTTGTTTGTATCTATTTTGCATGTGTTTGTAGATCTCAGATCAGTGTTTTAGAAAATCC  
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 144**

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY  
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP  
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15



## FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG  
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA  
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA  
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT  
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA  
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCAACGGCTGCATCAATGCCA  
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG  
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT  
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT  
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCCGAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT  
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA  
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT  
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT  
GAGAAGAACACATCAGGCACCTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTTTAGAGGTAG  
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC  
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT  
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT  
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

## **FIGURE 146**

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK  
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ  
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMVK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 157-171

**N-glycosylation sites.**

amino acids 98-102, 110-114

**Tyrosine kinase phosphorylation site.**

amino acids 76-83

**N-myristoylation sites.**

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

**Amidation site.**

amino acids 62-66

## **FIGURE 147**

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT  
CCTGATGATTTATAGACTCAAAGAAACTCATGTTTCAGAAGCTCTTCTCTTCTGGCCTCCTCT  
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG  
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCACTGAGTCTTAGTTTTTATTTTTTGAATTT  
CAACTTTCAGATTCAGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGCTGAGG  
TTTGGGGT

Downloaded from www.ashg.org

## **FIGURE 148**

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE  
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

bioRxiv preprint doi: <https://doi.org/10.1101/2021.08.11.456111>; this version posted August 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 149

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC  
TCCCGCGCGCCCCAACCTTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC  
CGGCCCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCTGTTTTACTCCTCCTTTTCATTGATA  
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT  
CCTCGGGACCGGCACTTGGAATCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCCAAACCTGGAGGAA  
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA  
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT  
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA  
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG  
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA  
TCACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGTGCCAGGATTTATGAAG  
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC  
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA  
TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG  
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA  
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA  
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTTCAGAAAAAGAACAAAAGAGAAAGAAA  
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT  
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT  
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA  
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTGAAGGATTCCACAAAA  
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC  
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT  
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA  
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAAAATGGCAAAGATCCAGGAGTCTTTCAA  
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAAATTTTTTGACCCAAGG  
GTTATTAGAAAGTGTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC  
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 150**

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG  
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK  
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE  
VLQKLISKEANNYEEDPNKPTSWTENQAGKIKEKVT PMAAIQDGLAKGENDET VSNLTLTNGLE  
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV  
SYLENLDEMIALQTKNLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD  
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHKDKGNKEDYDL SKMRDFINKQADAYVEKGILDK  
EEAEA IKRIYSSL

### **N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

### **Casein kinase II phosphorylation site:**

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-  
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-  
341, 369-372, 382-385, 386-389, 387-390

### **N-myristoylation sites:**

amino acids 143-148, 239-244

## FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA  
 GATGCTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC  
 ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC  
 GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATCCTGGGTGTCCAGGGTGGAAG  
 CCAGTGCCTGTTCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG  
 AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGCGGGACATGGGGCTCACC  
 TCCAGCTTCGAGTCGGCTGCCACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC  
 TGTGAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC  
 AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT  
 GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCACGCTGACTTAG  
 TGGGCACCTGACCCTTTGTCTTCTGGTTCCAGTTTGGATAAAATTTCTGAGATTTGGAGCTCAGT  
 CCACGGTCTCCCCACTGGATGGTGTCTACTGCTGTGGAACCTTGTAAAAACCATGTGGGGTAAA  
 CTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGCT  
 TAATGGTAAGTACAAGTGTACCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA  
 GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA  
 GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA  
 TGCTACCTTTCCCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGATGTCAGAA  
 GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA  
 GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGTCTATGACATATTGAGA  
 AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT  
 TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT  
 AGGTGATTTTTCTTTAATTTCTGTTAATTTATCTGTATTTCCCTAATTTTTCTACAATGAAGATGA  
 ATTCCTTGATAAAAAATAAGAAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA  
 TTGTCCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTTGAATCGAGCTCTGCTGCTCTGGTTGG  
 TTGTAGTAGTGATCAGGAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT  
 GTGGCTGGAATCTCTGGGTAAGGAACCTAAAGAACAATAATCATCTGGTAATTTCTTCTAGAAAG  
 GATCAGAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA  
 ATTTGTCCCCCTCAAATTCACATCCTTCTTGAATCTCAGTCTGTGAGTTTATTTGGAGATAAG  
 GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCATAT  
 GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA  
 AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC  
 ATCAGAAGCTTGAAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG  
 CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAAATTTCCGGCTGTTTTAA  
 GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT  
 CTCCTCGTGTTTACATTCTGTGTGTGTCCTTCCCACAATGTACCAAAGTTGTCTTTGTGACCAA  
 TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC  
 TACTTGAGCCCTCTCTCTGTGCCACCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG  
 AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT  
 GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT  
 AAGTTGCTCAGTTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

10063732

## **FIGURE 152**

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG  
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQ  
VRLTQLPENGWGNAPITDFYFQQCD

### **N-myristoylation sites.**

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

### **Interleukin-1 signature.**

amino acids 111-131

### **Interleukin-1 proteins.**

amino acids 8-29, 83-120, 95-134, 64-103

20250303 15:33:33



## FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC  
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG  
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC  
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA  
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT  
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTT  
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA  
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC  
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT  
GCCTGCATTTGACCCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA  
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG  
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA  
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAACCTGGTGTC  
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT  
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA  
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTATTTAT  
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT  
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

## **FIGURE 154**

MAALQKSVSSFLMGTLATSCLLLLLALLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL  
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS  
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-33

**N-glycosylation sites.**

amino acids 54-58, 68-72, 97-101

**N-myristoylation sites.**

amino acids 14-20, 82-88

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 10-21

## FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT  
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCA GTGCAGAGGGC  
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC  
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT  
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTTGCAT  
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG  
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG  
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT  
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCCGT  
TGCCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCGGGGCAACTCGGA  
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA  
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGCCCCGT  
GTGATGGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA  
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCTGTGAAGTGTGTCTGGAGCAG  
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG  
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCTATTTCTCTCAGGAAAGGTTTCAAA  
GTTCTGCCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG  
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT  
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT  
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCCAAATAAATAT  
CTTTATTTAAAAATGAAAAA

## **FIGURE 156**

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCP\$KGQDTSEELLRWSTVPVPPLEPA  
RPNRHPESCRASEDGPLNSRAISFWRYYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS  
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

20250315 15:50:04

## FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC  
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC  
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACAGTGTGCAACAGGGGACTATTCA  
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA  
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG  
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA  
GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG  
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA  
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA  
GTAGAAGTGAACCTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC  
TATCATCGGGTTTTCTCAGGTGTTTGTAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA  
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC  
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCAACAGGCGTCCCTTTCCCTCT  
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGGCCA  
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT  
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA  
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCAGAGTGAGGTGATCCTTGAAA  
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCACTGGCTTGCCACTCAAAGAAGGCA  
GCAGACAAAGTCGTCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA  
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCCTTTAACCTTTTCTGCA  
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA  
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT  
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG  
ATGGCTGCTGCTCCTTGTAG

## **FIGURE 158**

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPDGLRDLRVEPVTTTSVATGDYSILMNVS  
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP  
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH  
STIIGFSQVFEPHQKKQTRASVVIPTVDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGTGVPFPLDNNK  
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL  
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA  
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD  
GCCSL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-14

#### **Transmembrane domain:**

amino acids 290-309

#### **N-glycosylation sites.**

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283  
- 287

#### **cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 228 - 232 and 319 - 323

#### **Casein kinase II phosphorylation sites.**

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

#### **N-myristoylation site.**

amino acids 116-122

#### **Amidation site.**

amino acids 488-452

## **FIGURE 159**

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT  
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA  
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC  
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG  
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA  
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA  
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT  
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC  
ACTCAGCTGAAGAAG

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 160**

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN  
QRVSMERNIESRSTSPWNYTWTWDPNRYPSVVQAQCRNLGCINAQKEDISMNSVPIQQETLVV  
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

### **Signal sequence:**

amino acids 1-30

### **N-glycosylation site.**

amino acids 83-87

### **N-myristoylation sites.**

amino acids 106-111, 136-141



## FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG  
ACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGGATT  
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGGCAGCAC  
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGTTCTTGCTGTCT  
TGGCACTGGGCCGAAGCCAGTGGTCCTTTCTGAGAGAGGCTTGTGGGGCCTCAGGACGCTACC  
CACTGCTCTCCGGGCTCTCCTGCCGCTCTGGGACAGTGACATACTCTGCCTGCCTGGGACAT  
CGTGCTGCTCCGGGCCCCGTGCTGGCGCTACGCACCTGCAGACAGAGCTGGTGCTGAGGTGCC  
AGAAGGAGACCGACTGTGACCTGTCTGCGTGTGGCTGTCCACTTGGCCGTGCATGGGCACTGG  
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA  
TGCCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCCTACCTACTGCCCGCTGCGTCTGC  
TGGAGGTGCAAGTGCTGCTGCCCTTGTGCAGTTTGGTCAGTCTGTGGGCTCTGTGGTATATGAC  
TGCTTCGAGGCTGCCCTAGGAGTGAGGTACGAATCTGGTCTATACTCAGCCCAGGTACGAGAA  
GGAACCAACACACACAGCAGCTGCCTGCCCTGCCCTGGCTCAACGTGTACGAGATGGTGACA  
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTCGGCCTCTCCCTGTACTGGAATCAG  
GTCCAGGGCCCCCAAAACCCCGTGGCACAAAAACCTGACTGGACCGCAGATCATTACCTTGAA  
CCACACAGACCTGGTTCCCTGCCCTCTGTATTCAAGTGTGGCCTCTGGAACCTGACTCCGTTAGGA  
CGAACATCTGCCCTTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCGGCACTG  
CGACTGCTGACCCCTGCAGAGCTGGCTGCTGGACGCACCGTGTCTGCTGCCCGCAGAAGCGGCACT  
GTGCTGGCGGGCTCCGGGTGGGGACCCCTGCCAGCCACTGGTCCCACCGCTTTCTGGGAGAACG  
TCACTGTGGACAAGTTCTCGAGTTCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTCAAGTG  
AACAGCTCGGAGAAGCTGCAGTGCAGGAGTCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGA  
CGATGTGCTACTGTTGGAGACACGAGGCCCCAGGACAACAGATCCCTCTGTGCCTTGAACCCA  
GTGGCTGTACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGAGAGTACTTACTA  
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTG  
CCCCATGGACAAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCG  
CTGCGCTTTCCCTCATCTCTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAA  
CAGGACGTCCGCTCGGGGGCGGCCAGGGGCCGCGGCTCTGCTCTCTACTCAGCCGATGA  
CTCGGGTTTCGAGCGCTGGTGGGCGCCCTGGCGTCCGCCCTGTGCCAGCTGCCGCTGCGCGTGG  
CCGTAGACCTGTGGAGCCGTCTGAACTGAGCGCGCAGGGGCCCGTGGCTTGGTTTACGCGCAG  
CGCGCCAGACCTGCAGAGGGCGGCGTGGTGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCT  
GTGCAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCTTCC  
GCGCCTCGCTCAGCTGCGTGTCTGCCGACTTCTTGAGGGCCGGGCGCCCGCAGCTACGTGGGG  
GCCTGCTTCGACAGGCTGTCTCCACCCGACGCCGTACCCGCCCTTTTCCGACCGTGGCGCTCTT  
CACACTGCCCTCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCGTTCCG  
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAGCTACTTC  
CATCCCCGGGACTCCCGCGCCGGGACGCGGGTGGGACCAGGGGCGGGACCTGGGGCGGGGA  
CGGGACTTAAATAAAGGCAGACGCTGTTTTCTAAAAAAA

## **FIGURE 162**

MPVPWFLLSLALGRSPVVLRLERLVGPQDATHCSPGLSCRLWDSDIILCLPGDIVPAPGPFVLAPTHLQTELV  
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV  
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVVSADGDNVHLVLNV  
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN  
LWQAARLRLTLQSWLLDAPCSLPAAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ  
VNSSEKLQIQECLWADSLGPKDDVLLLETRGPQDNRLCALEPSGCTSLPSKASTRAARLGEYLLQDLQS  
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSAAAARG  
RAALLLYSADDGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP  
GAVALCSEWLQDGVSGPGAGHPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT  
LPSQLPDFLGALQQPRAPRSGRLQERAEQVSRAALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

### **Signal sequence:**

amino acids 1-20

### **Transmembrane domain.**

amino acids 453-475

### **N-glycosylation sites.**

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,  
334-337, 357-360, 391-394

### **Glycosaminoglycan attachment site.**

amino acids 583-586

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 552-555

### **N-myristoylation sites.**

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,  
692-697, 696-701, 700-705

## FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT  
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTTGA  
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA  
 AGACGTACCGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGC  
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT  
 GTCAGTGGGGAGGCGGTGAGCCACCAAGATGACTGACAGGTTGAGCTCTCTGCAGCACACTAC  
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA  
 CCCCCACGCCAATCCGTGACGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG  
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA  
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA  
 CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC  
 TACTCCTTCTCCGGAGCCTTCTCGTTCTCCATGGGCTTCTCGTTCGAGTACTCTGCTACCTGAG  
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT  
 TCCAGCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC  
 AGTCTGGCCCAGCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGAGCCCCGAGGAGC  
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC  
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCTGAG  
 GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATCTACGCCCCACA  
 GGCCATCTCTAAGGTCCAGCCTTCTCTATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT  
 CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCCTAAA  
 CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT  
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC  
 CCCTGGGGATTGTCACAGACAGAATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA  
 CCACAGTACCTAAAGGGCCAGTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC  
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC  
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCCCTGAGACCTCAGACCTG  
 GAGCAGCCACAGAACTGGATTCTTTTTAGAGGCCTGGCCCTGACTGTGCAGTGGGAGTCCCTG  
 AGGGGAATGGGAAGGCTTGGTGCTTCTCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA  
 ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC  
 AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA  
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCTGCTCGCTGAGCCCTG  
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC  
 CTAACACCATGGATTCAAAGTGCTCAGGGAAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC  
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTATTGTTCAAAGGTGGGAAGAGA  
 GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGTGGGCAGAACCAAGAACCTGC  
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA  
 TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTGAGCTTCATTCTCTGATAGAACAAGC  
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCT  
 ATCCTGAGAATGGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC  
 TGTACTGATGTCACAACCTTTGCAAGCTCTGCCTTGGGTTGAGCCATCTGGGCTCAAATTCAGC  
 CTCACCACTCACAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCCTC  
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG  
 TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA  
 AAAAAAA

## **FIGURE 164**

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW  
VAKKGCQRITRKSCNLTVETGNLTelyyARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCTIS  
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT  
EFLGTIMICVPTWAKESAPYMCRVKTLDPDRTWTYSFSGAFLESMGFLVAVLCYLSYRYVTKPPAP  
PNSLNVQRVLTFQPLRFIQEHVLI PVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT  
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPSSYAPQVTPEAQFPFYAPQAISKVQPSSY  
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSIAM  
EESQEAKSLHQPLGICTDRTSDPNVLHSGEETFPQYLKGQLPLLSSVQIEGHMPSLPLQPPSGPC  
SPSDQGPSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

### **Signal sequence.**

amino acids 1-17

### **Transmembrane domain.**

amino acids 233-250

### **N-glycosylation sites.**

amino acids 80-83, 87-90, 172-175

### **N-myristoylation sites.**

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

## FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATGG  
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG  
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGACGAAGAGTGCAGCAT  
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA  
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACTTTGGGCTGGA  
AGTGTGAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA  
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT  
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAACTGGAGGACTCTGTACCTGAAGAG  
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA  
ACTTGACCCTGTGCCTGAGCCCGAGGCATTGAGAGCTGATTGAGAGGATGGAGAAGGTGCTTTCT  
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT  
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA  
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA  
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT  
CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA  
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTTCCCTTCTAAGTA  
AAAAAAAAAAAAAAAAAAAA

## **FIGURE 166**

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPDCRFVN  
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPAETDFVCFEGGRD  
DFNSYNVEELLGSLELEDVPEESKKAEEVSHREKSPEESRGRELDVPEPEAFRADSEDGEGA  
FSESTEGLOGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLVPGSESRTGNSSPASVER  
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**N-glycosylation site.**

amino acids 294-298

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-34

**Tyrosine kinase phosphorylation site.**

amino acids 67-76

**N-myristoylation sites.**

amino acids 205-211, 225-231, 277-283

**Amidation site.**

amino acids 28-32

## FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAGCGC  
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC  
 AGTCCCCCAAACGCGCACCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGGGCACAGG  
 CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGGGCTCGGGC  
 GCGGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG  
 GGCAGAGGCCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG  
 TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCAGC  
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAGT  
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCA  
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG  
 ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG  
 TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG  
 ATGAACCTTCCTGCGGAAGTGAAAAGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGC  
 CTTGGGGGTCCTACCTTTACCAGTGGAAATGATGACAGGTGAACATGAAGCACAATTATATTTG  
 CAAGTATGAACCAGAGATTAATCCAACAGCCCCGTAGAAAAGCCTTATCTTACAAATCAACCAG  
 GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT  
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCT  
 GCATAAAAAGTAAAGGAAGAACAAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA  
 CCAGAAAAGAAAGTGGCATGGAAGTAATAAATACTCATTGACTTGGTTCCAGAATTTTGTAATTCT  
 GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAAATCACAAAGGATCTGC  
 AAGATGAAGTGAAGCTCCCCCTTGAGGCCAAATATTAAAGTAATTTTTATATGTCTATTATTTCA  
 TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA  
 ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA  
 TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTTACCTTTTACATAAGTTGTTATCTAGTCAATGTAA  
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGGTTTGATAAAA  
 ATGAAGTGTCTAATATTTATTTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG  
 AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAGTTTGT  
 TTTCTCGAAATAATTCATCTTTTACGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA  
 AATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT  
 TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTGTAGCTTAAATTAACAGATT  
 TTGTAATAATGTAACCTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG  
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA  
 GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAAGATGGTTGTTTCGGGGTTTGGG  
 ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG  
 ACTATATTAGTATACAAAGAGGTGATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG  
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACTACGAAATCGTGTGAAAATGGGTGG  
 AACCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATGGTGGCTCCTTT  
 CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT  
 CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 168**

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE  
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ  
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK  
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRITKTSNPQ  
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217